

Repeated-measures t -test

The t -test assesses whether the mean scores from two experimental conditions are statistically different from one another. A repeated-measures t -test (also known by other names such as the 'paired-samples' or 'related' t -test) is what you should use in situations when your design is within participants. In a within-participants design, participants contribute data for the dependent variable in ALL of the experimental conditions.

If you have a study design where different participants take part in your different experimental conditions, then you have a between-participants design. In this case you would need to use the independent t -test. There is a separate tutorial for this type of test.

The following example demonstrates what happens after you have created the data file. See the tutorial on Adding Variables for more information on how to create your own file.

Worked Example: doing a Repeated-Measures t -test in Jamovi

This tutorial will walk you through how to run and interpret a repeated-measures t -test in Jamovi. In this tutorial we will examine fictional data based on a study by Muncurlkin et al. (2021) that was inspired by the 'Hungry Judges' effect. This is the finding that hunger can influence judges' sentencing decisions, with studies showing that judges may be more lenient when sentencing after a meal break, and more severe before a meal break.

To test this effect, the researchers arranged a day of testing, in which participants were given ten vignettes (or short descriptions) which depicted fictional crimes committed by a fictional individual. For each vignette, participants scored how severe a sentence they would give to that crime if they were the judge from 1, least severe, to 10, most severe.

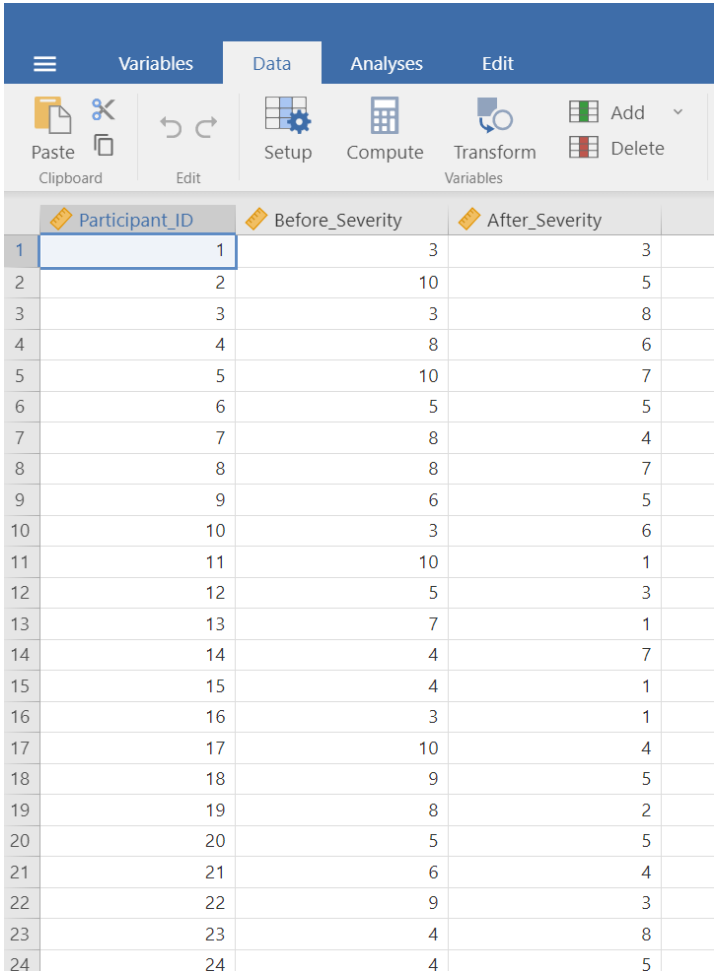
Crucially, participants saw five of the vignettes in the morning before a lunch break and five of the vignettes after a lunch break. The seriousness of the crimes depicted in the vignettes before and after lunch would need to be matched, and the order they were presented should be counterbalanced to prevent any order effects occurring.

As we are interested in the effect that a lunch break might have on sentencing severity, **lunch break** is our **independent variable (IV)**. It has two conditions, before the break and after the break. As participants see all the vignettes,

delivered both before and after lunch, this is a within-participants, also known as a paired-samples or repeated-measures, design.

As the variable we are measuring is the severity of the sentence the participants give to the crimes described in the vignettes, our **dependent variable (DV)** is **sentence severity**. This is measured by taking the mathematical average (or mean) of the severity scores for the five vignettes given before and after the lunch break. In this example, higher scores indicate more severe judgments.

This is what the data looks like in Jamovi.



	Participant_ID	Before_Severity	After_Severity
1	1	3	3
2	2	10	5
3	3	3	8
4	4	8	6
5	5	10	7
6	6	5	5
7	7	8	4
8	8	8	7
9	9	6	5
10	10	3	6
11	11	10	1
12	12	5	3
13	13	7	1
14	14	4	7
15	15	4	1
16	16	3	1
17	17	10	4
18	18	9	5
19	19	8	2
20	20	5	5
21	21	6	4
22	22	9	3
23	23	4	8
24	24	4	5

The different columns display the following data:

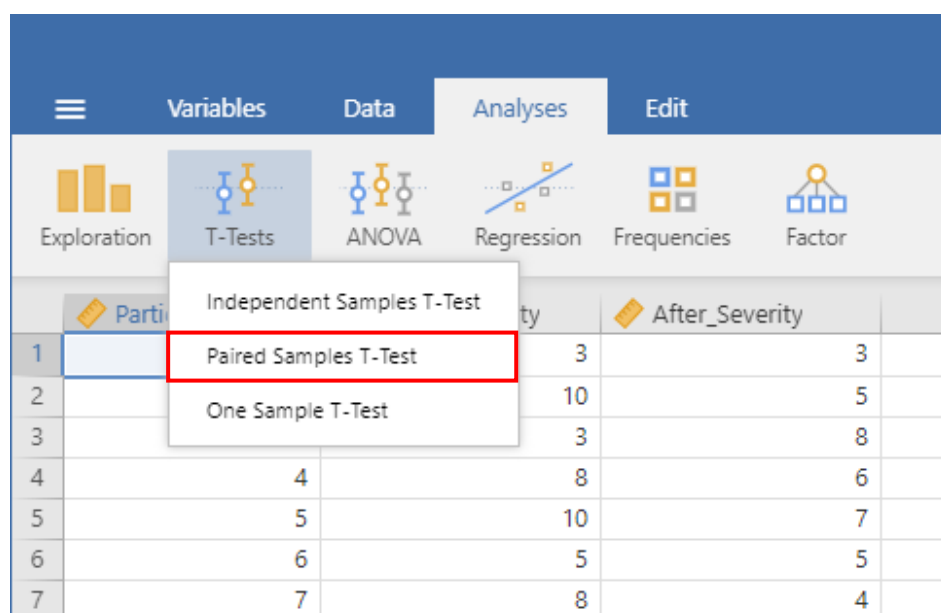
- **Participant_ID**: This just refers to the ID number assigned to the participants. We use these numbers as identifiers instead of participant names, as this allows us to keep participants' data together, while also keeping the participants anonymous. This is good practice in psychology, especially when collecting potentially sensitive data.

- **Before_Severity** : This column contains the mean severity scores for the five different vignettes seen before lunch (e.g., if a participant scored the five vignettes before lunch as 2, 2, 4, 4, 3, then the score shown for that participant in this column would be the mean of those data, which is 3). These are the DV scores for the first experimental condition.
- **After_Severity** : This column contains the mean severity scores for the five different vignettes seen after lunch. These are the DV scores for the second experimental condition.

Have you spotted the difference between this data file and the one we used for the independent t -test?

While the data file for an independent t -test contains a specific column that defines and codes the conditions of the independent variable (IV), this is not the case for a **repeated-measures t -test**. Instead, the within-participants IV is 'encoded' by the columns representing the two conditions under which the dependent variable was measured (e.g. before severity and after severity). Using this approach ensures that all the data from a single participant is on just one row in the data file.

To start the repeated-measures t -test analysis, click on Analyses, select T-Tests and click "Paired Samples T-test".



This brings up the Paired Samples T-test dialog box.

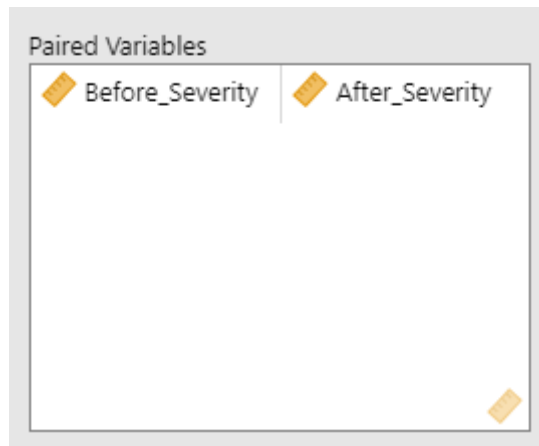
The screenshot shows the 'Paired Samples T-Test' configuration window in Jamovi. The top navigation bar includes 'Variables', 'Data', 'Analyses' (active), and 'Edit'. Below the navigation bar are icons for various statistical tests: Exploration, T-Tests, ANOVA, Regression, Frequencies, and Factor. The main window is titled 'Paired Samples T-Test' and features a search icon in the top right corner. On the left, a list of variables from the data file is shown: 'Participant_ID', 'Before_Severity', and 'After_Severity'. 'After_Severity' is currently selected. An arrow points from this list to a 'Paired Variables' box on the right, which is currently empty. Below the variable lists are several sections: 'Tests' (with 'Student's' selected and 'Bayes factor' and 'Wilcoxon rank' unchecked), 'Hypothesis' (with 'Measure 1 ≠ Measure 2' selected), and 'Missing values' (with 'Exclude cases analysis by analysis' selected). To the right of these are 'Additional Statistics' and 'Assumption Checks' sections, each with several unchecked options.

Here we can see all our variables from the data file displayed in the box on the left.

To tell Jamovi what we want to analyse we need to move our two variables that represent our DV to the correct box on the right.

First, **SELECT** Before_Severity in the left-hand box and **CLICK** on the arrow to move it to the right-hand box titled Paired Variables.

Then, **SELECT** After_Severity in the left-hand box and **CLICK** on the arrow to move it to the right-hand box titled Paired Variables.



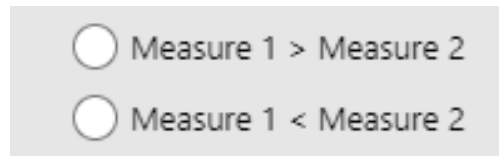
In the Paired Variables box (see images above), you can see that both conditions of your experiment now create one pair (Pair 1). In other words, they provide a pair of data points for Jamovi to compare (hence the paired samples terminology used by Jamovi to describe repeated measures designs). If you wanted to conduct more than one *t*-test at this point, you could insert more pairs of conditions into this box.

Before running the *t*-test in Jamovi, go to the **Hypothesis** section and make sure that you confirm whether you have a two-tailed (non-directional) or one-tailed (directional) hypothesis.

Two-tailed hypothesis: If your hypothesis is that there will be a difference between the conditions, but you do not specify which conditions you expect to have higher scores, then your hypothesis is **two-tailed**, as either condition resulting in significantly higher scores than the other would support your hypothesis. If this is the case, you would select this box below:

One-tailed hypothesis: If your hypothesis is that there will be a particular difference between the conditions (for example, that participants would be more severe in sentencing before lunch than after it), then your hypothesis is **one-tailed**, as only significantly higher scores from that condition over the other condition would support your hypothesis.

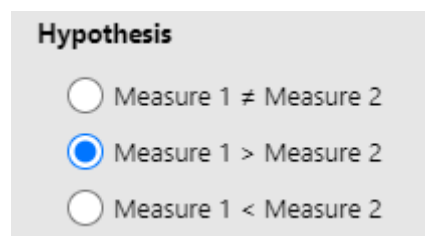
For a one-tailed hypothesis you need to select one of the options below, specifying the direction of your prediction (Measure 1 will have higher [$>$] scores than Measure 2, or Measure 1 will have lower [$<$] scores than Measure 2):



☐ Measure 1 > Measure 2

☐ Measure 1 < Measure 2

In the current example, our hypothesis is that sentences will be more severe before lunch. This is a one-tailed hypothesis, as we have predicted that scores will be higher before lunch (Measure 1) than after lunch (Measure 2). We therefore select this option:



Hypothesis

☐ Measure 1 $\neq$ Measure 2

☒ Measure 1 > Measure 2

☐ Measure 1 < Measure 2

Before running the t -test, we also want to make sure that we have selected:

- Descriptives

Once this has all been done your selection should mirror the image below, and the output will update on the right-hand side of the window.

Paired Samples T-Test

Participant_ID

Before_Severity

After_Severity

Before_Severity

After_Severity

☒ Student's
☐ Bayes factor

Prior 0.707

☐ Wilcoxon rank

☐ Mean difference
☐ Confidence interval 95 %
☐ Effect size
☐ Confidence interval 95 %
☒ Descriptives
☐ Descriptives plots

☐ Measure 1 ≠ Measure 2
☒ Measure 1 > Measure 2
☐ Measure 1 < Measure 2

Missing values

☒ Exclude cases analysis by analysis
☐ Exclude cases listwise

Additional Statistics

☐ Normality test
☐ Q-Q Plot

Results

Paired Samples T-Test

	statistic	df	p
Before_Severity After_Severity Student's t	5.36	106.00	< .001

Note: $H_a: \mu \text{ Measure 1} - \text{Measure 2} > 0$

Descriptives

	N	Mean	Median	SD	SE
Before_Severity	107	6.12	6.00	2.31	0.22
After_Severity	107	4.39	4.00	2.30	0.22

References

[1] The jamovi project (2022). *jamovi*. (Version 2.3) [Computer Software]. Retrieved from <https://www.jamovi.org>.

[2] R Core Team (2021). *R: A Language and environment for statistical computing*. (Version 4.1) [Computer software]. retrieved from MRAAN snapshot 2022-01-01).

The Output

So, what does the output show you? Let's look at the tables one at a time.

Group Descriptives:

We are going to look at the second and final box in our output first, as it provides us with descriptive statistics for our two conditions. It is always useful to inspect this box before you look at anything else, as it allows you to gain initial insight into the pattern of your data.

Descriptives					
	N	Mean	Median	SD	SE
Before_Severity	107	6.12	6.00	2.31	0.22
After_Severity	107	4.39	4.00	2.30	0.22

We are mainly interested in the **mean** and the **(SD) standard deviation** here. We can see from the two means that participants made more severe sentencing decisions before the lunch break (**mean = 6.12**) than after the lunch break (**mean = 4.39**). We can also see from the standard deviations (Before_Severity, **SD = 2.31**, After_Severity, **SD = 2.30**), that the scores in both conditions are similarly dispersed.

Note the value under **N** refers to the number of participants in each condition. Don't forget, though, that the same participants take part in both conditions, so in this case, we had **107 participants** participate in both conditions.

But how should you interpret the difference between the means? To find out whether this observed difference between the scores is statistically significant, you next need to look at the first table in our output, titled: Paired Samples T-Test.

Paired Samples T-test:

The first table in our output gives us our t -test statistics. We will go through all the columns and statistics that you will need to understand and write up your Paired samples t -test.

Paired Samples T-Test

Paired Samples T-Test

			statistic	df	p
Before_Severity	After_Severity	Student's t	5.36	106.00	< .001

Note. $H_a: \mu \text{ Measure 1} - \text{Measure 2} > 0$

Statistic (t-test statistic): This is the t -value calculated by the repeated measures t -test. This is an important statistic that you will need to report when writing up your findings.

It is an expression of the difference between the scores in your two experimental conditions. The larger the value of t , the more pronounced the difference between your conditions and the smaller the probability that this difference occurred by chance.

Df – Degrees of Freedom : It is an important statistic that needs to be reported, as its value directly impacts the significance of the t -statistic. In a repeated-measures t -test the value of df will be one less than the number of participants in the study (in this case there are 107 participants, so $df = 106$).

P (Significance): The significance columns give us the significance value (or p -value) for our test. If the value is smaller than .05, then the result is statistically significant. If it is larger, so statistically non-significant, then we reject our hypothesis in favour of the null hypothesis, which is that there is no difference between the two groups.

Writing up your t-test results

When writing up the results of your t-test you need to report whether the test was significant following this formula:

$$t(df) = t \text{ value}, p = p \text{ value}$$

...where you insert the relevant numbers from our table into the formula. For this example, we have found that the t-test is significant as the p-value is less than 0.05 ($p < .05$). This is reported as:

$$t(106) = 5.36, p < .001$$

Putting it all together

When interpreting and writing up your findings you need to use information from *both* the descriptive and the inferential statistics in your output. It doesn't matter which order you report these two types of statistics but always finish by interpreting your results meaningfully.

Step One: Describe the pattern of your data using the means and standard deviations from the first output table. In this case you could say something like:

Results showed participants made harsher sentencing decisions before lunch ($M = 6.12, SD = 2.31$) than after lunch ($M = 4.39, SD = 2.30$).

Step Two: Report the test you used, if it was one- or two-tailed, and whether or not the finding was *significant*:

A one-tailed repeated-measures t-test found this difference to be significant,
 $t(106) = 5.36, p < .001$.

Step Three: Finally, you need to put this information together to interpret and summarise what you have found in terms of your hypothesis. This should be written *in plain English*, for example:

This finding supported our hypothesis that sentencing severity would be affected by hunger.

What next?

Now you have learned how to carry out a repeated-measures t -test using Jamovi, why not try downloading the data file and see if you can produce the same output on your own?