

How to Perform McNemar's Test in R for Paired Data

Authored by
stats writer

March 10, 2026

RECOMMENDED CITATION

stats writer (2026). *How to Perform McNemar's Test in R for Paired Data*.

PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=135089>

McNemar's Test is a specialized non-parametric statistical procedure designed to analyze **paired categorical data**. Unlike standard chi-squared tests that assume independence between samples, this test is specifically tailored for scenarios where the same subjects are measured twice, typically before and after a specific intervention or treatment. By focusing on the changes that occur within subjects, researchers can determine whether a significant shift in **proportions** has taken place, providing deep insights into the effectiveness of the intervention being studied.

The primary utility of this test lies in its ability to evaluate **binary outcomes**--situations where the response can only be one of two possibilities, such as "Yes/No," "Success/Failure," or "Support/Do Not Support." In the context of clinical trials, marketing research, or social sciences, this method allows for a rigorous comparison of **marginal frequencies**. When we apply this test, we are essentially looking at the "discordant pairs"--those individuals whose status changed from the first observation to the second--while ignoring those whose status remained consistent.

In the **R programming language**, performing this analysis is highly efficient due to built-in functions that handle the complex underlying calculations. Statistical practitioners value **R** for its precision and the transparency of its diagnostic outputs. By using **R**, you can quickly transition from raw data entry to a comprehensive **statistical significance** analysis, ensuring that your conclusions are backed by robust mathematical evidence and adhering to the highest standards of academic and professional research.

Perform McNemar's Test in R

Understanding the Fundamentals of McNemar's Test

To effectively utilize McNemar's Test, one must first grasp the concept of **paired data**. This statistical method is applied when researchers are interested in the **proportions** of a population at two different time points or under two different conditions. Because the data points are not independent--meaning the second measurement depends on the individual's state during the first measurement--traditional methods like the Pearson's chi-squared test are inappropriate. Instead, this test specifically examines the internal consistency of the subjects over time.

The mathematical focus of this test is on the off-diagonal elements of a 2x2 contingency table. These elements represent the individuals who changed their response. For instance, if a person supported a policy before an event but opposed it afterward, or vice versa, they contribute to the test statistic. Those who stayed the same (concordant pairs) do not influence the final result regarding the change in proportions. This unique focus allows the test to be incredibly sensitive to shifts in public opinion or clinical status that might otherwise be obscured by the overall volume of the data.

When conducting research, the null hypothesis for this test typically states that the marginal probabilities for each outcome are the same across both conditions. Rejecting this hypothesis suggests that the intervention--whether it be a medication, a marketing campaign, or an educational seminar--had a statistically significant impact on the group. Understanding this foundation is crucial before moving into the technical implementation within the **R** environment, as it informs how we structure our **matrix** and interpret the resulting **p-value**.

Furthermore, the test assumes that the samples are representative of the population and that the pairs are independent of each other. While the two measurements for a single individual are related, the individuals themselves must be selected independently to ensure the validity of the results. This balance between **paired measurements** and independent subjects is what gives the test its power and precision in longitudinal studies or repeated measures designs across various scientific disciplines.

The Importance of Paired Categorical Data Analysis

Analyzing categorical data presents unique challenges compared to continuous data. In many real-world scenarios, we are not interested in the average change of a variable, but rather the frequency of a specific categorical shift. For example, in medical diagnostics, a researcher might compare the accuracy of two different screening tests on the same group of patients. Here, the data is inherently **paired**, and the goal is to see if one test consistently identifies a condition more frequently than the other, requiring a robust statistical framework.

The significance of using the correct test for **paired proportions** cannot be overstated. If a researcher were to use an independent samples test on paired data, they would likely overestimate the variance and fail to find a significant result that actually exists. This error in judgment could lead to the abandonment of a beneficial treatment or the failure to recognize a successful marketing strategy. By employing McNemar's Test, analysts account for the correlation within pairs, thereby increasing the **statistical power** of their findings.

Moreover, this approach is vital in behavioral sciences where pre-test and post-test designs are standard. Whether measuring the reduction of phobias after therapy or the adoption of a new technology after a training session, the focus is always on the transition between categories. The ability to quantify this transition allows for a clearer narrative of change, providing stakeholders with actionable data that goes beyond mere percentages and enters the realm of confirmed **statistical significance**.

Finally, as data becomes more complex, the ability to isolate specific variables through **paired analysis** becomes a competitive advantage for data scientists. It allows for the control of confounding variables that are inherent to the individual, such as age, gender, or baseline opinions, since the individual acts as their own control. This inherent control mechanism makes the

resulting **p-value** a much more reliable indicator of the effect of the variable being tested rather than external noise.

Practical Application: A Marketing Video Case Study

Suppose researchers want to know if a certain marketing video can change people's opinion of a particular law. They survey 100 people to find out if they do or do not support the law. Then, they show all 100 people the marketing video and survey them again once the video is over. This classic pre-and-post study design is the perfect candidate for McNemar's Test because it seeks to measure the impact of a single stimulus on a single group of individuals over time.

The following table shows the total number of people who supported the law both before and after viewing the video. This **contingency table** is the primary data structure required for our analysis. It categorizes the 100 participants into four distinct groups: those who supported the law both times, those who changed from support to non-support, those who changed from non-support to support, and those who never supported the law. By visualizing the data this way, we can immediately see where the shifts are occurring.

Before Marketing Video		
After Marketing Video	Support	Do not support
Support	30	40
Do not Support	12	18

To determine if there was a statistically significant difference in the proportion of people who supported the law before and after viewing the video, we look at the discordant cells. In this case, 40 people moved from "Do Not Support" to "Support," while 12 people moved from "Support" to "Do Not Support." The test will evaluate whether this imbalance--40 versus 12--is large enough to suggest that the video had a real effect, rather than the change being due to random chance or sampling error.

This scenario highlights the importance of the "Before" and "After" columns. If the number of people switching from support to opposition was roughly equal to the number switching from opposition to support, the overall **proportions** would remain the same, and the test would result in a high **p-value**. However, because there is a visible disparity in our example, we anticipate that the test will reveal a significant impact caused by the marketing intervention.

Step 1: Data Preparation and Matrix Construction in R

The first practical step in performing the test is to create the dataset in a format that **R** can process. Typically, this is done by constructing a **matrix**. In **R**, a matrix is a collection of data elements

arranged in a two-dimensional rectangular layout. For our purposes, we need a 2x2 matrix where the rows represent the "After" state and the columns represent the "Before" state. Proper labeling of these dimensions is essential for maintaining clarity during the analysis and interpretation phases.

By using the `matrix()` function, we can input our observed counts directly. It is important to ensure that the order of the data matches the intended structure of the **contingency table**. In the code snippet below, we use the `c()` function to combine our counts and the `nrow` argument to specify the dimensions. Additionally, the `dimnames` argument is utilized to provide descriptive labels for our rows and columns, which significantly improves the readability of the output when the **matrix** is printed to the console.

```
#create data
```

```
data <- matrix(c(30, 12, 40, 18), nrow = 2,  
dimnames = list("After Video" = c("Support", "Do Not Support"),  
"Before Video" = c("Support", "Do Not Support")))
```

```
#view data
```

```
data
```

```
Before Video
```

```
After Video Support Do Not Support
```

```
Support 30 40
```

```
Do Not Support 12 18
```

Once the **matrix** is created, it is always a best practice to view the data by calling the object name. This allows you to verify that the 30, 12, 40, and 18 counts are in their correct relative positions. Specifically, you want to ensure that the discordant pairs (12 and 40) are correctly identified in the off-diagonal cells. If the data is misaligned at this stage, the subsequent McNemar's Test will produce an incorrect **p-value**, leading to a flawed interpretation of the experiment's results.

In more complex workflows, you might start with a **data frame** containing raw individual responses. In such cases, you can use the `table()` function in **R** to automatically generate the required 2x2 structure. Regardless of the method used to reach the **matrix** stage, the goal remains the same: to produce a clean, well-organized summary of the observed shifts in the categorical variables, ready for the application of the statistical test.

Step 2: Executing the Statistical Procedure in R

With the data properly structured, we can now proceed to perform the actual analysis. In **R**, the primary tool for this is the `mcnemar.test()` function. This function is part of the standard **stats**

package, meaning no external library installations are required for basic usage. The syntax is straightforward but allows for specific arguments that can alter the way the test statistic is calculated, providing the user with flexibility depending on their specific dataset requirements.

The standard syntax for the function is `mcnemar.test(x, y = NULL, correct = TRUE)`. Here, the variable `x` represents our **contingency table** or **matrix**. The `y` argument is generally ignored if `x` is already a matrix. The `correct` argument is particularly important, as it determines whether a **continuity correction** will be applied. Understanding when to set this to `TRUE` or `FALSE` is a critical skill for any data scientist or statistician working with **paired categorical data**.

x: This should be either a two-dimensional contingency table in **matrix** form, or a factor object representing the first set of observations.

y: This is a factor object representing the second set of observations; it is ignored if `x` is already provided as a matrix.

correct: This logical value indicates whether to apply the **continuity correction** when computing the test statistic. By default, **R** sets this to `TRUE`.

The execution of this function will generate a wealth of information, including the **chi-squared** test statistic, the **degrees of freedom**, and the all-important **p-value**. The degrees of freedom for a standard 2x2 McNemar's Test is always 1, as the test is fundamentally comparing the difference between two counts. By carefully reviewing these outputs, you can determine if the observed change in **proportions** is large enough to be considered meaningful in a scientific context.

The Role and Impact of the Continuity Correction

One of the more technical aspects of McNemar's Test in **R** is the application of the continuity correction, often referred to as Yates's correction. This adjustment is designed to compensate for the fact that we are using a continuous distribution (the **chi-squared** distribution) to approximate a discrete distribution. Without this correction, the test can sometimes be overly aggressive, yielding a **p-value** that is slightly too small and increasing the risk of a Type I error.

In general, a **continuity correction** should be applied when some counts in the table are small. A common rule of thumb in statistics is to apply this correction when any of the cell counts--specifically the discordant cell counts--are less than 5 or 10. However, in modern computational statistics, many researchers prefer to run the test both ways to see how much the correction influences the final result. If both versions lead to the same conclusion, the researcher can be more confident in the stability of their findings.

We will perform the analysis both with and without the correction to illustrate the mathematical differences in the output. Note how the **chi-squared** value decreases slightly when the correction is applied, which in turn leads to a slightly higher **p-value**. This conservative approach is often

preferred in peer-reviewed research to ensure that the findings are not an artifact of the mathematical approximation used during the testing process.

#Perform McNemar's Test with continuity correction

```
mcnemar.test(data)
```

McNemar's Chi-squared test with continuity correction

```
data: data
```

```
McNemar's chi-squared = 14.019, df = 1, p-value = 0.000181
```

#Perform McNemar's Test without continuity correction

```
mcnemar.test(data, correct=FALSE)
```

McNemar's Chi-squared test

```
data: data
```

```
McNemar's chi-squared = 15.077, df = 1, p-value = 0.0001032
```

As observed in the results above, the choice to apply the **continuity correction** does change the raw numbers, but the fundamental conclusion remains the same. In professional settings, if your discordant pairs are sufficiently large (e.g., both are greater than 20), the impact of the correction becomes negligible. However, for smaller sample sizes, the `correct = TRUE` setting is a vital safeguard that ensures the integrity of your statistical significance claims.

Deciphering the Test Output and Making Informed Decisions

The final and most important stage of the process is the interpretation of the **p-value**. In the context of McNemar's Test, the p-value tells us the probability of observing a difference in **proportions** as large as ours (or larger) assuming that the null hypothesis is true. If this probability is very low, we conclude that the change we observed is unlikely to have happened by chance alone, and we reject the null hypothesis in favor of the alternative.

In our marketing video example, the **p-values** produced (0.000181 and 0.0001032) are both significantly lower than the standard alpha level of 0.05. This indicates a very high level of **statistical significance**. Therefore, we can confidently reject the null hypothesis and conclude that the proportion of people who supported the law before and after watching the marketing video was significantly different. The video appears to have been highly effective in shifting the participants' opinions toward supporting the law.

When reporting these results, it is essential to include the test statistic, the **degrees of freedom**, and the **p-value**. This transparency allows others to verify your work and understand the strength

of the evidence. Furthermore, while the test tells us that a significant change occurred, it is also helpful to look back at the raw data to describe the direction of the change. In our case, the shift was overwhelmingly positive, as 40 people moved to "Support" while only 12 moved away from it.

In summary, McNemar's Test is an indispensable tool for any researcher working with **paired categorical data** in **R**. By carefully constructing your **matrix**, choosing the appropriate correction settings, and rigorously interpreting the **p-value**, you can transform raw survey or clinical data into compelling, scientifically-backed conclusions. Whether you are evaluating a marketing strategy or a new medical protocol, this test provides the mathematical rigor needed to make truly informed decisions.

ARABPSYCHOLOGY.COM