# Encyclopedia of Research Design

## Pretest–Posttest Design

Contributors: Bethany A. Bell
Editors: Neil J. Salkind
Book Title: Encyclopedia of Research Design
Chapter Title: "Pretest–Posttest Design"
Pub. Date: 2010
Access Date: January 14, 2014
Publishing Company: SAGE Publications, Inc.
City: Thousand Oaks
Print ISBN: 9781412961271
Online ISBN: 9781412961288
DOI: http://dx.doi.org/10.4135/9781412961288.n331
Print pages: 1087-1092

http://dx.doi.org/10.4135/9781412961288.n331
The basic premise behind the pretest–posttest design involves obtaining a pretest measure of the outcome of interest prior to administering some treatment, followed by a posttest on the same measure after treatment occurs. Pretest–posttest designs are employed in both experimental and quasi-experimental research and can be used with or without control groups. For example, quasi-experimental pretest–posttest designs may or may not include control groups, whereas experimental pretest–posttest designs must include control groups. Furthermore, despite the versatility of the pretest–posttest designs, in general, they still have limitations, including threats to internal validity. Although such threats are of particular concern for quasi-experimental pretest–posttest designs, experimental pretest–posttest designs also contain threats to internal validity.

# Types of Pretest–Posttest Designs without Control Groups

# One-Group Pretest-Posttest Design

In the simplest pretest–posttest design, researchers gather data about some outcome through a single pretest, administer a treatment, and then gather posttest data on the same measure. This design is typically represented as follows:

$$O_1 \quad X \quad O_2$$

where O
*1*

represents the pretest, X represents some treatment, and O
*2*

represents the posttest.

Including a pretest measure is an improvement over the posttest-only research design; however, this design is still relatively weak in terms of internal validity. Although this design allows researchers to examine some outcome of interest prior to some treatment (O
*1*

), it does not eliminate the possibility that O
*2*

might have occurred regardless of the treatment. For example, threats to internal validity, such as maturation, history, and testing, could be responsible for any observed difference between the pretest and posttest. Also, the longer the time lapse between the pretest and posttest, the **[p. 1087 ↓ ]** harder it is to rule out alternative explanations for any observed differences.

When the outcome of interest is continuous, data obtained from the one-group pretest–posttest design can be analyzed with the dependent-means *t* test (also sometimes called the correlated-means *t* test or the paired-difference *t* test), which tests *H*
*0*

: μ
*D*

= 0. When the outcome of interest is categorical and has only two levels, one-group pretest–posttest data can be analyzed with McNemar's chi-square to test the null hypothesis that the distribution of responses is the same across time periods. However, when the outcome of interest is categorical and has more than two levels, McNemar's chi-square cannot be used. Instead, data can be analyzed through Mantel-Haenszel methods.

# One-Group Pretest-Posttest Design Using a Double Pretest

The one-group pretest–posttest design can be improved upon by adding a second pretest prior to treatment administration:

$$O_1 \quad O_2 \quad X \quad O_3,$$

where O
*1*

and O
*2*

represent the two pretests, X represents some treatment, and O
*3*

represents the posttest.

Adding a second pretest to the traditional one-group pretest–posttest design can help reduce maturation and regression to the mean threats as plausible explanations for any observed differences. For example, instead of comparing O
*3*

only to O
*1*

or O
*2*

, any observed difference between O
*2*

and O

*3*

can also be compared to any differences between O
*1*

and O
*2*

. If the difference between O
*2*

and O
*3*

is larger than the difference between O
*1*

and O
*2*

, then the observed change is less likely solely due to maturation.

When the outcome of interest is continuous, data obtained from the one-group pretest–posttest design using a double pretest can be analyzed through a within-subject design analysis of variance (ANOVA) (also sometimes referred to as repeated measures ANOVA) with appropriate contrasts to test the null hypothesis of interest, $(O_{1A}, O_{1B}) \text{ X } (O_{2A}, O_{2B})$

. When the outcome of interest is categorical, data from this design can be analyzed with conditional logistic regression (also referred to as subject-specific models) to test the null hypothesis that the distribution of responses is the same across time periods.

# One-Group Pretest-Posttest Design Using a Nonequivalent Dependent Variable

The simple one-group pretest–posttest design also can be improved by including a nonequivalent dependent variable. For example, instead of obtaining pretest and posttest data on the outcome of interest, with this design, the researcher obtains pretest and posttest data on the outcome of interest (A) as well as on some other outcome (B) that is similar to the outcome of interest, but is not expected to change based on the treatment.

$$(O_{1A}, O_{1B})\ X\ (O_{2A}, O_{2B})$$

In this design, O
*1A*

represents the pretest on the outcome of interest, O
*1B*

represents the pretest on the nonequivalent outcome, X represents some treatment, O
*2A*

represents the posttest on the outcome of interest, and O2B represents the posttest on the nonequivalent outcome. Including a non-equivalent dependent variable helps researchers assess any naturally occurring trends in the data, separate from the treatment effect. If the posttest difference for the outcome of interest (O
*2A*

–O
*1A*

) is larger than the posttest difference for the secondary outcome (O
*2A*

–O
*1B*

), the less likely the change is solely a result of maturation.

When the outcome of interest is continuous, data obtained from the one-group pretest-post-test design using a nonequivalent dependent variable can be analyzed through a within-subject ANOVA design with appropriate contrast statements to test the null hypothesis of interest,

$$H_0: \mu_{\text{postA}} - \mu_{\text{preA}} = \mu_{\text{postB}} - \mu_{\text{preB}}$$

When the outcome of interest is categorical, data from this design can be analyzed with conditional logistic regression to test the null hypothesis that the distribution of responses is the same across time periods and across the two dependent variables (i.e., outcome of interest and the nonequivalent dependent variable).

# Removed-Treatment Design

This design is also an improvement to the one-group pretest–posttest design. By adding multiple **[p. 1088 ↓ ]** posttests and the removal of the treatment to the one-group pretest–posttest design, the removed-treatment design allows better control over the internal validity of a study. That is, by examining the outcome of interest after the treatment has been stopped or removed, researchers have another piece of information that allows for multiple comparisons and helps inform their conclusions. This design is typically represented as

$$\mathrm{O}_1\ \mathrm{X}\ \ \mathrm{O}_2\ \mathrm{O}_3\ \mathrm{X}\ \mathrm{O}_4,$$

where O
*1*

represents the pretest on some outcome of interest, X represents some treatment, O
*2*

represents the first posttest on the outcome of interest, O
*3*

represents the second posttest on the same outcome of interest, X represents treatment removal, and O
*4*

represents the third posttest on the outcome of interest.

With this design, if a treatment effect exists, the difference between O
*2*

and O
*1*

should be greater than the difference between O
*4*

and O
*3*

. Also, examining changes between O
*3*

and O
*2*

allows the researcher to investigate possible trend effects after treatment. However, this theoretical pattern of change is often difficult to observe unless a treatment effect dissipates quickly upon treatment removal (which is often not the case in the social and behavioral sciences). Furthermore, because removal of certain treatments might be considered unethical, considerations need to be discussed before employing this pretest–posttest design.

When the outcome of interest is continuous, data obtained from the removed-treatment design can be analyzed through a within-subject ANOVA with appropriate contrast statements to test the null hypothesis of interest,

$$O_1 \; X \; O_2 \; X \; O_3 \; X \; O_4$$

. When the outcome of interest is categorical, data from this design can be analyzed with conditional logistic regression to test the null hypothesis that the distribution of responses is the same across time periods.

# Repeated-Treatment Design

Building upon the removed-treatment design, the repeated-treatment design includes multiple posttest observations, treatment removal, and treatment reintroduction. In the following notation, O
*1*

represents the pretest on some outcome of interest, X represents some treatment, O
*2*

represents the first posttest on the outcome of interest, X—represents treatment removal, O
*3*

represents the second posttest on the same outcome of interest, and O
*4*

represents the third posttest on the outcome of interest.

$$O_1 \; X \; O_2$$

$$O_1 \quad O_2,$$

If a treatment effect is present, in this design, O
*2*

should differ from O
*1*

and from O
*3*

; however, the difference between O
*3*

and O
*2*

should be opposite or less than the difference between O
*2*

and O
*1*

. Similarly, O
*4*

should differ from O
*3*

, and the difference between O
*4*

and O
*3*

should be similar to the difference between O
*2*

and O
*1*

. Although the inclusion of a second treatment period adds rigor to this design, it is not without limitations. As with the removed-treatment design, this design is best when the treatment effect will truly dissipate upon removal of the treatment and when it is ethically acceptable to remove the treatment. Furthermore, by reintroducing the treatment, participants are more likely to become aware of the treatment and researchers' expectations, thus increasing the potential of contamination.

When the outcome of interest is continuous, data obtained from the repeated-treatment design can be analyzed through a within-subject design ANOVA with appropriate contrast statements to test the null hypotheses of interest,

$$H_0 : \mu_{\text{post1}} - \mu_{\text{pre}} = \mu_{\text{post3}} - \mu_{\text{post2}}$$

. When the outcome of interest is categorical, data from this design can be analyzed with conditional logistic regression to test the null hypothesis that the distribution of responses is the same across time periods.

# Types of Pretest–Posttest Designs with Control Groups

# Two-Group Pretest-Posttest Design

Similar to the one-group pretest–posttest design, this design includes a nontreated control group and is often represented as follows:

$$\mathrm{O}_1 \ \mathrm{X} \ \mathrm{O}_2$$
$$\mathrm{O}_1 \quad \mathrm{O}_2,$$

**[p. 1089 ↓ ]**

where O

*1*

represents the pretest, X represents some treatment, and O
*2*

represents the posttest. When randomization is employed, the two-group pretest–posttest design becomes the classic experimental design.

Whether the two-group pretest–posttest design is used in experimental or quasi-experimental research, including an untreated control group reduces threats to internal validity and allows for within-group and between-group comparisons. For example, although selection bias is still a concern when this design is used in quasi-experimental research, use of a pretest and a comparison group allows researchers to examine the nature and extent of selection bias by comparing the treatment and control groups before the treatment is administered. Noting any pretest differences between the groups allows for stronger inferences to be made after the treatment is administered. Also, although threats such as maturation and history still might exist with the two-group pretest–posttest design, the effects of these threats should be the same for both groups, thus adding more support for a treatment effect when observed posttest differences exist between the groups.

When the outcome of interest is continuous, data obtained from the two-group pretest–posttest design can be analyzed with a one-between one-within ANOVA design (also referred to as a factorial repeated measures ANOVA) to test

$$O_1\ O_2\ X\ O_3$$

$$O_1\ O_2\quad O_3$$

. When the outcome of interest is categorical, two-group pretest–posttest data can be analyzed with generalized estimating equations (GEEs) to test the null-hypothesis that the distribution of responses is the same across time periods and between groups.

# Two-Group Pretest-Posttest Design with a Double Pretest

Improving upon the two-group pretest–posttest design, the two-group pretest–posttest design with a double pretest includes administering two pretests prior to treatment followed by one posttest measure after the treatment period.

$$\begin{array}{ccc} O_1 & X & O_2 \\ O_1 & & O_2 \\ & X & O_2 \\ & & O_2 \end{array}$$

In this design, O
*1*

represents the first pretest, O
*2*

represents the second pretest, X represents some treatment, and O
*3*

represents the posttest. One advantage of the two-group pretest–posttest design with a double pretest over the two-group pretest–posttest design is the ability to examine trends and selection bias in the treatment and control groups prior to treatment administration. If the treatment has an effect, the change between O
*3*

and O
*2*

for the treatment group should be different from the observed change during the same time period for the control group. This design also reduces threats to internal validity. By including the untreated control group, threats such as maturation, history, and regression to the mean would occur in both groups, thus, in essence, balancing the groups in terms of threats. Therefore, observed pretest–posttest treatment group differences are more likely to represent treatment effects.

When the outcome of interest is continuous, data obtained from the two-group pretest–posttest design with a double pretest can be analyzed with a one-between one-within ANOVA design with appropriate contrast statements to test the null hypothesis of interest,

$$H_0 : \mu_{\text{post\_Tx}} - \mu_{\text{pre2-\_Tx}} = \mu_{\text{post\_C}} - \mu_{\text{pre2\_C}}$$

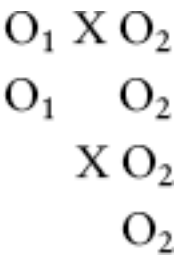
. When the outcome of interest is categorical, data from this design can be analyzed with GEEs with appropriate contrast statements to test the null hypothesis,

$$H_0 : \mu_{\text{post\_Tx}} - \mu_{\text{pre2-\_Tx}} = \mu_{\text{post\_C}} - \mu_{\text{pre2\_C}}.$$

.

# Four-Group Design with Pretest-Posttest and Posttest-Only Groups

Also known as the Solomon four-group design, this design is a combination of the two-group pretest–posttest design and the two-group posttest-only design. Pretest measures are obtained from the two pretest–posttest groups, treatment is administered to one of the pretest–posttest groups and to one of the posttest-only groups, and posttest measures are obtained from all four groups.




**[p. 1090 ↓ ]**

In this design, O
*1*

represents the pretest, X represents some treatment, and O
*2*

represents the posttest.

Although this design is a bit more complicated than the simple two-group pretest–posttest design, the inclusion of the two posttest-only groups allows researchers to investigate possible testing threats to internal validity. For example, if O
*2*

for the treatment pretest–posttest group is similar to O
*2*

for the treatment posttest-only group, then testing effects are likely not present. Similarly, as with other multiple-group pretest–posttest designs, this design allows for both within-group and between-group comparisons, including examination of possible selection bias when this design is used in quasi-experimental research. With this design, maturation and history threats to internal validity are also diminished. If a treatment effect is present, then the difference between O
*2*

and O
*1*

should be the same for the treatment and control pretest–posttest groups, and O
*2*

for the two treatment groups should be similar to one another and different from O
*2*

for the two control groups.

When the outcome of interest is continuous, data obtained from the Solomon four-group design can be analyzed with two different statistical procedures. First, a one-between one-within ANOVA design can be used to test the null hypothesis
$\mu_{\text{post\_Tx}} - \mu_{\text{pre\_Tx}} = \mu_{\text{post\_C}} - \mu_{\text{pre\_;C}}$

and an independent means *t* test can be used to test the null hypothesis
$\mu_{\text{post\_Tx2}} - \mu_{\text{post\_C2}}$

. When the outcome of interest is categorical, data can be analyzed with GEEs to test analogous null hypotheses:
$H_0: \mu_{\text{post\_Tx}} - \mu_{\text{pre\_Tx}} = \mu_{\text{post\_C}} - \mu_{\text{pre\_C}}$ and $H_0 : \mu_{\text{post\_Tx2}} - \mu_{\text{post\_C2}}$

and
$H_0: \mu_{\text{post\_Tx}} - \mu_{\text{pre\_Tx}} = \mu_{\text{post\_C}} - \mu_{\text{pre\_C}}$ and $H_0 : \mu_{\text{post\_Tx2}} - \mu_{\text{post\_C2}}$

.

# Switching Replications Design

The switching replications design refers to the process of obtaining pretest data from two groups, administering some treatment to one group, obtaining a second assessment measure from both groups, administering the same treatment to the second group, and obtaining a third assessment measure from both groups. This design is typically denoted as

$$O_1 \; X \; O_2 \quad O_3$$
$$O_1 \quad O_2 \; X \; O_3,$$

where O
*1*

refers to the pretest assessment for both groups, X represents some treatment, O
*2*

represents the first posttest for the first group and the second pretest for the second group, and O
*3*

represents a second posttest for the first group and the first posttest for the second group.

As with the other pretest–posttest designs that use a control group, if participants are not randomly assigned to the groups, selection bias is a likely threat. However, because pretest measures are obtained from both groups, potential group differences can be examined before the first treatment is administered. Furthermore, administering the treatment to both groups at different times helps improve the internal validity of the study. For example, if some trend were responsible for any observed difference between O
*2*

and O
*1*

for the first group, then one also would expect to see the same level of change during this time period in the second group. This same principle applies to differences between O
*3*

and O
*2*

. Thus, if a treatment effect exists, the difference between O
*2*

and O
*1*

for the first group should be different from the difference between O
*2*

and O
*1*

for the second group, yet similar to the difference between O
*3*

and O
*2*

for the second group. However, as with the four-group design with pretest–posttest and posttest-only groups, the timing difference in treatment administration is an important issue that needs to be acknowledged when employing this design. The treatment administration to the second group can never be exactly the same as the treatment that is administered to the first group.

When the outcome of interest is continuous, data obtained from the switching replications design can be analyzed with a one-between one-within ANOVA design with appropriate contrast statements to test the null hypotheses of interest,

$$H_0 : \mu_{\text{post1\_Tx}} - \mu_{\text{pre\_Tx}} = \mu_{\text{pre2\_C}} - \mu_{\text{pre1\_C}}$$

and

$$H_0 : \mu_{\text{post2\_Tx}} - \mu_{\text{post1\_Tx}} = \mu_{\text{post\_C}} - \mu_{\text{pre2\_C}}$$

. When the outcome of interest is categorical, data from this design can be analyzed using GEEs and analogous contrasts as those above, to test the null hypotheses

$$H_0 : \mu_{\text{post1\_Tx}} - \mu_{\text{pre\_Tx}} = \mu_{\text{pre2\_C}} - \mu_{\text{pre1\_C}}$$

and

$$H_0 : \mu_{\text{post2\_Tx}} - \mu_{\text{post1\_Tx}} = \mu_{\text{post\_C}} - \mu_{\text{pre2\_C}}$$

.

# Reversed-Treatment Pretest-Posttest Control Group Design

With this design, one group receives the treatment of interest, expected to affect the outcome of **[p. 1091 ↓ ]** interest in one direction, and the other group receives an opposite treatment, expected to affect the outcome of interest in the opposite direction. Diagrammed as

$$\mathrm{O}_1 \ \mathrm{X}_+ \ \mathrm{O}_2$$
$$\mathrm{O}_1 \ \mathrm{X}_- \ \mathrm{O}_2,$$

O
*1*

represents the pretest, X+ represents the treatment of interest, X
–

represents the treatment expected to produce opposite effects, and O
*2*

represents the posttest.

The basic premise behind this design is that if the difference between O
*2*

and O
*1*

for the group administered X
+

is in one direction and the difference between O
*2*

and O
*1*

for the group administered X
–

is in the opposite direction, a statistical interaction, suggesting a treatment effect, should be present. However, as with the other pretest–posttest designs that use a control group, if participants are not randomly assigned to the groups, selection bias is a likely threat and should be examined. Also, as with pretest–posttest designs that include treatment removal or delayed treatment administration, ethical issues often surround use of the reversed-treatment pretest–posttest control group design. Finally, conceptual difficulties are inherent in such a design as researchers must identify treatments that are expected to produce opposite effects on the outcome variable.

When the outcome of interest is continuous, data obtained from the switching replications design can be analyzed with a one-between one-within ANOVA design to test the null hypothesis of interest,

$$H_0 : \mu_{\text{post}} - \mu_{\text{pre2}} = \mu_{\text{pre2}} - \mu_{\text{pre1}}$$

. When the outcome of interest is categorical, data can be analyzed with GEEs to test the null hypothesis that the distribution of responses is the same across time periods and between groups.

Bethany A. Bell

http://dx.doi.org/10.4135/9781412961288.n331
*See also*

- Experimental Design
- General Linear Model
- Logistic Regression
- Quasi-Experimental Design
- Repeated Measures Design
- Threats to Validity
- Within-Subjects Design

Further Readings

Brogan, D. R., & and Kutner, M. H. Comparative analyses of pretest–posttest research design. American Statistician, (1980). vol. 34, pp. 229–232.

Feldman, H. A., and McKinlay, S. M. Cohort versus cross-sectional design in large field trials: Precision, sample size, and a unifying model. Statistics in Medicine, (1994). vol. 13, pp. 61–78.

Shadish, W. R., Cook, T. D., Campbell, D. T. (2002). Experimental and quasi-experimental designs for generalized causal inference. Boston: Houghton Mifflin.

Stokes, M. E., Davis, C. S., Koch, G. G. (2006). Categorical data analysis using the SAS system. Cary, NC: SAS Institute.