The SAGE Encyclopedia of Educational Research, Measurement, and Evaluation

Repeated Measures Designs

Contributors: Markus Brauer
Edited by: Bruce B. Frey
Book Title: The SAGE Encyclopedia of Educational Research, Measurement, and Evaluation
Chapter Title: "Repeated Measures Designs"
Pub. Date: 2018
Access Date: February 26, 2018
Publishing Company: SAGE Publications, Inc.
City: Thousand Oaks,
Print ISBN: 9781506326153
Online ISBN: 9781506326139
DOI: http://dx.doi.org/10.4135/9781506326139.n586
Print pages: 1408-1409

The defining characteristic of repeated measures designs is the fact that independent units—usually participants—are “crossed with” at least one of the independent variables; that is, each unit provides at least one data point for each level of one or more independent variables. In other words, in repeated measures designs, at least one of the independent variables varies “within units” and is thus referred to as a within-unit variable (e.g., within-subjects variable). In the most general sense, repeated measures designs are characterized by data that are clustered by participants (or other units) and are thus nonindependent. Repeated measures designs are different from purely between-subjects designs, in which participants are said to be “nested under” one or more independent variables.

In the simplest repeated measures design, each participant provides one data point for each of the two levels of a dichotomous independent variable. Common repeated measures designs are studies in which participants’ responses are collected twice (e.g., at the beginning and at the end of the school year) or in which each participant is exposed to multiple types of stimuli (e.g., each student evaluates one structured and one unstructured task). In more complex repeated measures designs, independent units are crossed with more than one independent variable or are crossed with some independent variables and nested under others. It is also possible for the within-subjects variable to have more than two levels (e.g., students’ performance is measured 5 times during the academic year).

Statistical Power and Internal Validity

Compared to purely between-subjects designs, repeated measures designs usually have greater statistical power. This is due to the fact that more data points are obtained with the same number of participants and that individual differences are accounted for and therefore do not contribute to the error term of the inferential test. Repeated measures designs frequently have lower internal validity, in that there might be alternative explanations for the observed differences between experimental conditions. Many of the threats to internal validity can be eliminated; however, one can include practice trials before the actual study to avoid learning effects. One can keep the task short or maintain a high level of motivation to do well on the task (e.g., by rewarding participants for good performance) to avoid fatigue effects. Finally, one can space out the measurement moments or include a distractor task between them to avoid carry-over effects from the first experimental condition to the second.

The best way to increase internal validity in a repeated measures design is to counterbalance the order of conditions. Half of the (randomly chosen) participants first do Condition 1 and then do Condition 2 of the independent within-subjects variable, whereas the other half of the participants proceeds in the inverse order. Statistical power is generally increased if order is subsequently included as a predictor in the statistical analyses. The analysis is then a mixed-models analysis of variance with one within-subjects variable (treatment) and one between-subjects variable (order). Depending on the data analysis software the researcher is using, it may be necessary to “center” the order variable (i.e., to recode it into −.5 and +.5 or into −1 and +1) to obtain the treatment effect averaged across order conditions.

In certain pretest–posttest designs, statistical power can be increased by treating the pretest as a covariate (sometimes called analysis of covariance approach or regression adjustment) rather than treating pretest and posttest as two levels of a within-subjects variable (sometimes called repeated measures approach or change score analysis). As noted by G. J. P. van Breukelen, the more powerful pretest-as-covariate approach can be used only if certain conditions are satisfied: (a) There is one (and only one) dichotomous within-subject variable, and one of the two levels is clearly a pretest or a baseline measure, (b) there is at least one between-subjects variable, and (c) the assignment to the levels of the between-subjects variables is either random or determined by participants' pretest score. The pretest-as-covariate approach consists of regressing the posttest on both the pretest and the between-subjects variables.

Advanced Techniques for Complex Designs

It is possible to statistically control for covariates in repeated measures designs. When the covariate varies between subjects (e.g., an individual difference measure), it suffices to add it to the regression model (like order). When the covariate varies within subjects (e.g., mood assessed at each measurement moment), it is sometimes called a time-varying covariate. According to Charles M. Judd, David A. Kenny, and Gary H. McClelland, the appropriate regression model is then $(Y_2 - Y_1) = b_0 + b_1 (Z_2 - Z_1) + b_2 ((Z_1 + Z_Z)/2) - C$, where $(Y_2 - Y_1)$ is the outcome difference, $(Z_2 - Z_1)$ is the covariate difference, and $((Z_1 + Z_2)/2) - C$ is the mean-centered covariate average. The inclusion of the last term is not absolutely necessary, but without it, one makes the (often unreasonable) assumption that the covariate-outcome relationship is the same in both experimental conditions. In the aforementioned equation, the coefficient b_0 tests the (within-subject) treatment effect, statistically controlling for the time-varying covariate.

It is also possible to examine mediation in repeated measures designs. By definition, the mediator has to vary within subjects. Mediation is tested with the same regression equation as shown earlier, the only difference being that Z_1 and Z_2now refer to the two mediator scores (one per experimental condition). The coefficient b_0 tests for the (within-subject) treatment effect, statistically controlling for the mediator (this effect is referred to as "Path *a*" in many relevant texts on mediation). The coefficient b_1 tests for the effect of the mediator on the outcome variable (usually referred to as "Path *b*").

In certain repeated measures designs, participants provide multiple responses for each level of the independent within-subjects variable. Sometimes all participants provide responses to the same targets or materials (e.g., there are 10 structured and 10 unstructured tasks, and all students evaluate the same set of 20 tasks) and sometimes each participant reacts to the participant's own unique set of targets or materials (e.g., each student is asked to nominate and judge 10 same-sex and 10 different-sex friends; each student evaluates a different set of 20 individuals). These types of studies are best analyzed with linear mixed-effects models. Note that these two designs require both a by-subject random intercept and a by-subject random slope, but that the former design—all students evaluate the same set of items—requires in addition a by-item random intercept, according to Charles M. Judd, Jacob Westfall, and David A. Kenny.

See alsoGeneralized Linear Mixed Models; Mediation Analysis; Mixed Model Analysis of Variance; Random Assignment; Regression Discontinuity Analysis

Markus Brauer
http://dx.doi.org/10.4135/9781506326139.n586
10.4135/9781506326139.n586

Further Readings

Brauer, M., & Curtin, J. J. (in press). Linear mixed-effects models and the analysis of nonindependent data: A unified framework to analyze categorical and continuous independent variables that vary within subjects and/or within items. Psychological Methods.

Judd, C. M., Kenny, D., & McClelland, G. H. (2001). Estimating and testing mediation and moderation in within-subject designs. Psychological Methods, 6, 115–134.

Judd, C. M., Westfall, J., & Kenny, D. (2017). Experiments with more than one random factor: Designs, analytic models, and statistical power. Annual Review of Psychology. doi:http://dx.doi.org/10.1146/annurev-psych-122414-033702

Montoya, A. K., & Hayes, A. F. (2016). Two-condition within-participant statistical mediation analysis: A path-analytic framework. Psychological Methods, 22, 6–27. doi:http://dx.doi.org/10.1037/met0000086

Van Breukelen, G. J. P. (2006). ANCOVA versus change from baseline had more power in randomized studies and more bias in nonrandomized studies. Journal of Clinical Epidemiology, 59, 920–925.