Why we prefer repeated measures designs over between-groups designs?

OR

Why Use Repeated Measures Designs? OR Advantages of repeated measure design

Because of the following advantages, repeated measures designs are often preferred over between-groups designs. The use of repeated measures, where the same subjects are tested under a number of conditions, has numerous practical and statistical benefits.

- Primary advantage is that it eliminates almost all "individual differences" problems.
- No individual differences between-groups, since there is only one group.
- Differences between treatments can be separated from individual differences, so no confounding.
- Individual differences are the same across all treatments, and can be removed, making treatment differences easier to see, thus increasing "power."
- The total number of individual differences is much smaller than for between-groups designs, so "chance" variability is greatly reduced, thus increasing "power."
- **More statistical power**: Repeated measures designs can be very powerful because they control for factors that cause variability between subjects.
- **Fewer subjects**: A repeated measures design can use fewer subjects to detect a desired effect size. Further sample size reductions are possible because each subject is involved with multiple treatments. For example, if an independent groups design requires 20 subjects per experimental group, a repeated measures design may only require 20 total.
- **Quicker and cheaper**: Fewer subjects need to be recruited, trained, and compensated to complete an entire experiment.
- **Assess an effect over time:** Repeated measures designs can track effect overtime, such as the learning curve for a task. In this situation, it's often better to measure the same subject at multiple times rather than different subjects at one point in time for each.

Furthermore, RMD provides more efficient estimates, RMD study changes in participant's behaviors across time. Repeated measures designs increase sensitivity by reducing error variance that would be due to having different participants in the various conditions (i.e. it increases statistical power). Each subject serves as own control so that the variability between subjects gets isolated. Analysis can focus more precisely on treatment effects. Repeated measures designs should only be used when carry-over effects are unlikely.

Disadvantages of repeated measure design

Major disadvantage of RMD is that Time related effects and order effects may influence the scores and thus confound the results.

Repeated Measures (Montgomery)

In experimental work in the social and behavioral sciences and some aspects of engineering the physical sciences, and business, the experimental units are frequently people. Because of differences in experience, training, or background, the differences in the responses of different people to the same treatment may be very large in some experimental situations. Unless it is controlled, this variability between people would become part of the experimental error, and in some cases, it would significantly inflate the error mean square, making it more difficult to detect real differences between treatments. In many repeated measures the experimental units are not necessarily people; they could be different stores in a marketing study, or plants, or experimental animals, and so on. We typically think of these experimental units as **subjects**.

It is possible to control this variability between people or "projects" by using a design in which each of the ***v*** treatments is used on each person. Such a design is called a **repeated measures design**. Here, we give a brief introduction to repeated measures experiments with a single factor. Suppose that an experiment involves ***v*** treatments and every treatment is to be used exactly once on each of **n** subjects. The data would appear as in **Table 1**. Note that the observation y_{ij} represents the response of subject *j* to treatment *i* and that only **n** subjects are used. The model that we use for this design is:

$$Y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij} \qquad \ldots\ldots\ldots\ldots\ldots (1)$$

Table 1. Data for a single factor Repeated Measures Designs

Treatments	**Subjects**				**Treatments Totals**
	1	**2**	**…**	**n**	
1	y_{11}	y_{12}	…	y_{1n}	$\mathbf{y_{1.}}$
2	y_{21}	y_{22}	…	y_{2n}	$\mathbf{y_{2.}}$
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
v	y_{v1}	y_{v1}	…	y_{vn}	$\mathbf{y_{2.}}$
Subject Totals	$\mathbf{y_{.1}}$	$\mathbf{y_{.2}}$	**…**	$\mathbf{y_{.n}}$	$\mathbf{y_{..}}$

Where τ_i is the effect of the *i*th treatment and β_j is a parameter associated with the *j*th subject. We assume that treatments are fixed (so $\sum_{i=1}^{v} \tau_i = 0$) and that the subjects employed are a random sample of subjects from some larger population of potential subjects. Thus, the subjects collectively represent a random effect, so we assume that the mean of β_j is zero and that the variance of β_j is σ^2_β. Because the term β_j is common to all *v* measurements on the same subject, the covariance between y_{ij} and $y_{\iota' J}$ is not, in general, zero. It is customary to assume that the covariance between y_{ij} and $y_{\iota' J}$ is constant across all treatments and subjects.

Principles of Analysis

- Repeated-measures analysis is a generalization of paired 't' test. The only difference is that measurements are made on the same individuals. These are likely to be correlated, and analysis must take such correlations into account.
- The total variability is partitioned into between subjects and within subjects.
- Next the within subjects variability is partitioned into explained by treatment and residual (unexplained) variability.
  Consider an analysis of variance partitioning of the total sum of squares, say

$$\sum_{i=1}^{v}\sum_{j=1}^{n}(y_{ij}-\bar{y}_{..})^2 = v\sum_{j=1}^{n}(\bar{y}_{.j}-\bar{y}_{..})^2 + \sum_{i=1}^{v}\sum_{j=1}^{n}(y_{ij}-\bar{y}_{.j})^2 \quad \ldots\ldots\ldots (1)$$

We may view the first term on the right-hand side of Equation (1) as a sum of squares that results from differences between subjects and the second term as a sum of squares of differences within subjects. That is,

$$\mathbf{SS}_T = \mathbf{SS}_{\text{Between Subjects}} + \mathbf{SS}_{\text{Within Subjects}}$$

The sums of squares $SS_{\text{Between Subjects}}$ and $SS_{\text{Within Subjects}}$ are statistically independent, with degrees of freedom:

$$\boldsymbol{vn - 1 = (n - 1) + n(v - 1)}$$

The differences within subjects depend on both differences in treatment effects and uncontrolled variability (noise or error). Therefore, we may decompose the sum of squares resulting from differences within subjects as follows:

$$\sum_{i=1}^{v}\sum_{j=1}^{n}(y_{ij}-\bar{y}_{.j})^2 = \mathrm{n}\sum_{i=1}^{v}(\bar{y}_{i.}-\bar{y}_{..})^2 + \sum_{i=1}^{v}\sum_{j=1}^{n}(y_{ij}-\bar{y}_{i.}-\bar{y}_{.j}-\bar{y}_{..})^2 \quad \ldots\ldots\ldots (2)$$

The first term on the right-hand side of Equation (2) measures the contribution of the difference between treatment means to $SS_{\text{Within Subjects}}$, and the second term is the residual variation due to error. Both components of $SS_{\text{Within Subjects}}$ are independent. Thus,

$$\mathbf{SS}_{\text{Within Subjects}} = \mathbf{SS}_{\text{Treatment}} + \mathbf{SS}_{\text{E}}$$

with the degrees of freedom given by

$$\boldsymbol{n(v - 1) = (v - 1) + (v - 1)(n-1)} \text{ respectively.}$$

To test the hypothesis of no treatment effect, that is,

$$\boldsymbol{H_0: \tau_1 = \tau_2 = \ldots = \tau_v = 0}$$

$$\boldsymbol{H_1:} \textbf{ At least one } \boldsymbol{\tau_i \neq 0}$$

we would use the ratio

$$\boldsymbol{F_0 = \frac{SS_{Treatment}/(v-1)}{SS_E/(v-1)(n-1)} = \frac{MS_{Treatments}}{MS_E}} \qquad \ldots\ldots\ldots (3)$$

If the model errors are normally distributed, then under the null hypothesis, H0: $\tau_i = 0$, the statistic F_0 follows an F distribution. The null hypothesis would be rejected if $F_0 > F_{\alpha, (v\text{-}1, (v\text{-}1)(n\text{-}1))}$

ANNOVA Table

S.O.V	**d.f**	**SS**	**MS**	**F_0**
B/W Subject	n-1	$\frac{1}{v}\sum_{j=1}^{n} y_{.j}^2$ - C.F		
Within Subject	n(*v*-1)	$\sum_{i=1}^{v}\sum_{j=1}^{n} y_{ij}^2$ - $\frac{1}{v}\sum_{j=1}^{n} y_{.j}^2$		
Treatment	*v*-1	$\frac{1}{n}\sum_{i=1}^{v} y_{i.}^2$ - C.F	SS_{Treat}/df	MS_{Treat}/ MS_E
Error	(n-1)(*v*-1)	Within Subj SS – Treat SS	SS_E/df	
Total	n*v*-1	$\sum_{i=1}^{v}\sum_{j=1}^{n} y_{ij}^2$ - C.F		

Practical Example

Three different Pinot Noir wines were evaluated by a panel of eight judges. The judges are considered a random panel of all possible judges. The wines are evaluated on a 100-point scale. The wines were presented in random order to each judge, and the following results obtained. Analyze the data from this experiment. Is there a difference in wine quality? Analyze the residuals and comment on model adequacy.

Judge	Wine			**Treat Totals**
	1	2	3	
1	85	88	93	**266**
2	90	89	94	**273**
3	88	90	98	**276**
4	91	93	96	**280**
5	92	92	95	**279**
6	89	90	95	**274**
7	90	91	97	**278**
8	91	89	98	**278**
Subject Totals	**716**	**722**	**766**	**2204**

C.F = G^2/nv = $(2204)^2/24$ = 202400.67

Total SS = $\sum_{i=1}^{v}\sum_{j=1}^{n} y_{ij}^2$ - C.F

= 202664 – 202400.67 = 263.33

B/W Subject SS = $\frac{1}{v}\sum_{j=1}^{n} y_{.j}^2$ - C.F

= $\frac{1}{3}[(266)^2 + (273)^2 + \ldots + (278)^2]$ – C.F

= $\frac{1}{3}(607346)$ – 202400.67 = 48

Within Subject SS = $\sum_{i=1}^{v}\sum_{j=1}^{n} y_{ij}^2$ - $\frac{1}{v}\sum_{j=1}^{n} y_{.j}^2$

= 202664-202448.67 = 215.33

Treat SS = $\frac{1}{n}\sum_{i=1}^{v} y_{i.}^2$ - C.F

= $\frac{1}{8}[(716)^2 + (722)^2 + (766)^2]$ – C.F

= 202587 – 202400.67 = 186.33

ANNOVA Table

S.O.V	**d.f**	**SS**	**MS**	**F_0**	**F_t**
B/W Subject	n-1 = 7	48.00			
Within Subject	n(*v*-1) = 16	215.33			
Treatment	*v*-1 = 2	186.33	93.165	44.98**	3.74
Error	(n-1)(*v*-1) = 14	29.00	2.0714		
Total	n*v*-1 = 23	263.33			

****