Letter to the Editor

A design by any other name….

Michael F.W. Festing
Understanding Animal Research

Abstract

The randomised block design is used mainly to increase the power of an experiment by matching experimental units receiving different treatments. It is widely used in many disciplines where it is known by at least eight different names. One of these, "within subjects" is confusing because the term subjects has to be defined to fit the design and another "repeated measures", is confusing because it is also used for an entirely different design. Failure to standardise nomenclature can cause confusion when teaching the principles of experimental design.

The randomised block design

The randomised block design (RBD) is a two-way design first developed for agricultural field experiments ([Fisher 1960](#)). It has one factor which is often random or possibly fixed but of little interest. This factor is variously called a *block*, a *subject*, a *trial*, a *replicate* or even an *experiment* (by some biologists).   The RBD also has one fixed effect factor, often called the *treatment* or named according to the type of factor, such as "diet" or "drug". The purpose of a RBD is mainly to increase the power of the experiment by controlling some sources of variability or by taking account of some natural structure among the experimental units. Blocking may also increase the generality of the results by comparing units under slightly different conditions and it is a convenient way of breaking an experiment up into more manageable parts. Assuming a quantitative dependent variable, a RBD is usually analysed by a two-way analysis of variance without interaction (assuming no within-block replication) fitting the model $y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij}$ where τ_i is the effect of treatment i: i = 1, 2,..., k, $\Sigma\tau_i = 0$, β_j is the effect of block j: j = 1, 2, ...,b, $\Sigma\beta j = 0$, and the ε_{ij} are assumed to be independent, normally distributed random variables with a mean of zero and constant variance σ^2 ([Snedecor and Cochran 1980](#)).

The naming of a design

If a "design" is defined in terms of a mathematical model, then *cross-over, change-over, within-subject, matched samples, related subjects, correlated samples, dependent samples* and *repeated-measures* (in one incarnation) designs are all also randomised block designs. They only differ in whether the blocks are animate (a subject) or inanimate (a block, trial, replicate or experiment) and whether replication within the block is in space, such as plots in a field or eyes in an ophthalmology experiment, or in time such as when a subject is given different treatments serially, assuming no carry-over effect. If there were thought to be a systematic trend within the blocks, all of these could be converted to a Latin square design by adding another random factor, as suggested by several authors (Cox and Reid 2000; Mead 1988).

Within and between subjects designs

The description of one-way completely randomised designs as "between subjects" and two-way designs as "within-subjects" that is widely used in behavioural studies (Howell 1999; Roberts and Russo 1999) is confusing because the definition of a subject has to be altered to fit the design. Thus a study in which twins are given either a placebo or a drug (a two-way design) is called a "within-subject" design even though the comparison is between the twins who could each legitimately be called a "subject". In contrast Cox and Reid (2000) describe "Typical ways of forming blocks….batches of material produced on one machine where several similar machines are producing nominally the same product…". Obviously, the machines are the blocks. But this means that if the experiment is done using only a single machine (a completely randomised design), it would be classified as a "between subjects" design. In this case the definition of "subject" has to be changed to a machine for a period of time (the experimental unit) to fit the situation, which is inconsistent.

The problem with repeated measures designs

Even more confusing is the use of the term "repeated-measures". Does it really have the same mathematical model as the randomised block design? Yes, according to Montgomery (1984) who states in his section on repeated-measures designs that "Readers … will recognise the analysis of variance for a single-factor design with repeated measures as equivalent to the analysis for a randomized complete block design, with subjects considered to be the blocks." Yes, also according to Snedecor and Cochran (1980) who discuss it in the same chapter as the randomised block design. Similarly, Howell (1999) gives an example of a "repeated-measures" design where people are assessed before and after training, with people being a random factor. But "no" according to Kaps and Lamberson (2009) who give a numerical example of a "repeated-measures" design investigating the effect of three treatments on milk fat yield in cattle with four cattle per group, fat being measured weekly for six weeks. They recognise that each fat measurement is not a separate experimental unit and fit an entirely different model with no blocking factor. The purpose was to estimate a treatment x time interaction, i.e. whether the change in butter fat levels was different among different treatments, rather than to increase statistical power. They do note, however, that "..'change-over' designs can be considered repeated measures designs, but they differ in that two or more treatments are assigned to each experimental subject."

Maxwell and Delaney (1989) suggest three ways in which the term "repeated-measures" is used. The first is as a randomised block with the same behavioural measure being used as the dependent variable in each different treatment condition with the aim being to increase power. The second is when "scores on each of *a* different tests are collected for each subject." The aim here being to compare the means of the *a* scores. The third situation is when some

aspect of behaviour may be measured at two or more times in the same subject but without a change of treatment in what might more appropriately be called a longitudinal design. This latter use of the design is controversial because, as pointed out by Mead (1988), measurements at different times are variables rather than experimental units and time is not a variable which can be randomised. He suggests the alternative of fitting derived variables such as the mean for each subject, the slope of a regression line or a multivariate analysis. Many investigators assume that if they are measuring some character repeatedly in a subject, it must be a repeated measures design.

Conclusions

It is confusing when one design can have several different names or when one name can be used for more than one design and the term "subject" has to be defined to fit the design. Because most of the names are more descriptive than RBD, particularly when applied to live subjects, it is unlikely that they will be abandoned. A solution might be:

1. When using one of these alternative names teachers could explain that the design is derived from RBD agricultural field experiments, but has found wide application in several other disciplines, under different names. The same mathematical model is fitted in each case. Students do not have to learn a different statistical analysis for each one.

2. It could be explained that the term "repeated measures" can be used to mean a RBD, but it may also be used for a design where each subject is measured repeatedly without different treatments being applied in what might be called a longitudinal design. In this case the purpose is generally not to increase power but to estimate change over time in each subject and extra care has to be taken to identify the experimental unit and ensure that the statistical analysis, using a different mathematical model, is correct.

3. Students could be told that although the classification of experiments into between and within subjects can be helpful, what is meant by a "subject" may not always be entirely clear.

4. Writers of textbooks might consider dealing with all the names for a RBD in a single chapter, rather than repeating the same statistical analysis in different chapters.

Acknowledgment

References

Cox, D. R. and Reid, N. (2000), *The theory of the design of experiments*, Boca Raton, Florida: Chapman and Hall/CRC Press.

Fisher, R. A. (1960), *The design of experiments*, New York: Hafner Publishing Company, Inc.

Howell, D. C. (1999), *Fundamental Statistics for the Behavioral Sciences*, Pacific Grove, London, New York: Duxbury Press.

Kaps, M. and Lamberson, W. (2009), *Biostatistics for animal science*, Wallingford: CAB International.

Maxwell, S. E. and Delaney, H. D. (1989), *Designing experiments and analyzing data*, Belmont, California: Wadsworth Publishing Company.

Mead, R. (1988), *The design of experiments*, Cambridge, New York: Cambridge University Press.

Montgomery, D. C. (1984), *Design and Analysis of Experiments*, New York: John Wiley & Sons, Inc.

Roberts, M. J. and Russo, R. (1999), *A student's guide to the analysis of variance*, London: Routledge.

Snedecor, G. W. and Cochran, W. G. (1980), *Statistical methods*, Ames, Iowa: Iowa State University Press.

Michael F.W. Festing,
Understanding Animal Research
25 Shaftsbury Avenue
London, UK.
Michaelfesting@aol.com

Volume 18 (2010) | Archive | Index | Data Archive | Resources | Editorial Board | Guidelines for Authors | Guidelines for Data Contributors | Home Page | Contact JSE | ASA Publications

4