

More Than Your Heart Desires...

...An Exploration of Blocking

Teacher Notes

Carolyn Doetsch, Ocean Lakes High School, Virginia Beach, VA
cdoetsch@worldnet.att.net

Peter Flanagan-Hyde, Phoenix Country Day School, Phoenix, AZ
pflanaga@pcds.org

Mary Harrison, Salem High School, Virginia Beach, VA
johne2@series2000.com

Josh Tabor, Wilson High School, Hacienda Heights, CA
jtabor@hlpusd.k12.ca.us

Chuck Tiberio, Wellesley High School, Wellesley, MA
tiberio@tiac.net

NCSSM Statistics Leadership Institute
July, 2000

1. Introduction

An experiment is used to determine the influence of one or more treatments on some desired response. The response of each individual in an experiment will vary, and thus there will be variation in the outcome of the experiment depending on who the actual subjects are. This comes from two sources: response to the treatment and response to other factors (anticipated or unanticipated) such as gender, age, weight, etc. To reduce variation due to anticipated factors, experimenters will block for those factors. To reduce variation due to unanticipated factors, they will use randomization.

In a blocked design, the subjects of the experiment are divided according to some characteristic that is believed to create variation in the response. A common example is gender: men and women often react differently to drugs in a clinical trial. Blocking on gender would mean that the group is first divided into men and women, and then each of the gender groups is randomly divided into treatment groups. This guarantees that there will be the same number of men in each treatment group and the same number of women in each treatment group, whereas complete randomization would only mean that *on average* there is the same number. Without blocking, some of the randomized treatment groups would have different numbers of men in each treatment group and different numbers of women in each treatment group, and if gender does influence the response, the results from these groups would vary more than groups which had been blocked.

Consider an extreme example to illustrate this point:

Suppose two drugs (A and B) are compared in a trial. All men, using either drug, improve by 20 points, and all women, using either drug, improve by 10 points. Clearly these drugs are equally effective, and work for both men and women. In a trial with 10 subjects in each treatment group, a blocked design will have 5 men and 5 women receiving each treatment. The average response in every trial will then be 15 for both drugs.

However, if the experiment is conducted with a completely randomized design, then some groups may have more men than women, or vice versa. Possibly 6 men and 4 women will be assigned to group A, with 4 men and 6 women assigned to group B. The average response would be 16 for drug A versus 14 for drug B, implying that drug A is better. In repeated trials, the means for each drug would average 15, but there would be variation from trial to trial.

However, blocking will reduce the variability in the means of the sample responses only if the factor chosen for the blocking is responsible for some of the variation in the response. In the example above, blocking by gender does reduce the variation because men and women reacted differently to the drug.

The reduction in variability because of blocking is greatest when there are as many blocks as possible. The ultimate goal is to have the number of subjects in each block equal to the number of treatments in the experiment.

The purpose of randomization is to reduce the variation due to unanticipated factors. Randomization will make it equally likely for a treatment group to get an individual who will have a particularly good or a particularly bad response, so that at least on average these good and bad responses are evenly distributed between the two treatments. Thus a completely randomized design, which relies only on randomization to form the treatment groups, may have more variation than a randomized block design.

The purpose of the classroom activity that accompanies this document is:

- to provide practice in the mechanics of blocking and randomization,
- to provide insight into why blocking is effective in reducing variation,
- to investigate why blocking works for some factors and not others.

2. The Data Set

This data set consists of human subjects involved in a medical experiment. For each subject 8 characteristics are recorded:

1. Gender (Coded with 0 = male, 1 = female)
2. Age (to the nearest year)
3. Exercise Level (Coded with 0 = none, 1 = occasional, 2 = frequent)

4. Initial Cholesterol
5. Reading after treatment with drug A
6. Reading after treatment with drug B
7. Improvement using A (Initial reading minus reading after taking drug A)
8. Improvement using B (Initial reading minus reading after taking drug B)

The data were constructed to illustrate the principles of blocking. Gender and exercise level were assigned so each category would have the same number of subjects. Age is determined by randomly assigning an integer from 30 to 60 inclusive to the subjects. Initial cholesterol, reading after A, and reading after B are determined using the formulas below. The response to drugs A and B depend on initial cholesterol and exercise level and include a term to create some random variation.

$$Init = \lfloor 300 - 10 \cdot Gen - 20 \cdot Exer + 10 \cdot RandNorm$$

$$A = \left\lfloor 15 + Init + 12 \cdot Exer - \frac{Init}{6} + 7.5 \cdot RandNorm$$

$$B = \left\lfloor 15 + Init + 15 \cdot Exer - \frac{Init}{5} + 4.5 \cdot RandNorm$$

Note: RandNorm is a normally distributed random variable with mean 0 and standard deviation 1.

3. Classroom Procedure

The first consideration is to what extent calculators will be used in this activity. Calculators can be used either for simple computation with group placement done by paper and pencil, or data can be entered into the calculator's lists and the process somewhat automated. For your convenience, detailed descriptions of these procedures for the TI-83 are available.

As you work through the classroom activity, here are some suggestions, possible answers, and potential pitfalls.

General comments:

- The scenario describes a fairly realistic example of how an HMO might decide on a particular drug for treating high cholesterol.
- Students may wonder why all 50 volunteers are not used in the experiment. Common practice is to get equal sized groups in each of the categories of interest (therefore the subjects chosen are not a random sample). In addition researchers try to use the minimum number of subjects to reduce costs.
- Make certain that students correctly fill in the improvement columns (noting that almost all values should be positive).

- Remember that positive values represent a decrease in cholesterol level, hence a higher value represents a better performance.

I. Completely Randomized Design

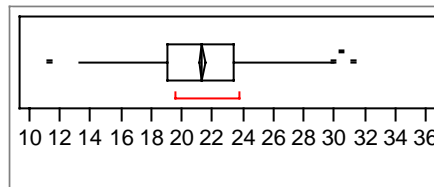
Make certain that the students sample without replacement (i.e., all 12 random integers must be different) and that the students use the improvement values(initial – final).

Using data from all of your students, you may want to construct 2 boxplots, one for the mean improvement in A and the other for mean improvement in B. The plots should show that drug B outperforms drug A but there is significant overlap (typically 20-25% of the time, the improvement using A will be greater than the improvement using B). Thus you may not be completely confident deciding that drug B is better.

Here are boxplots that illustrate the pattern you should see. These pictures were generated using the software JMP-IN 4 by randomizing the subjects 1000 times. The boxplots for your class will be somewhat different since you have a much smaller number of randomizations.

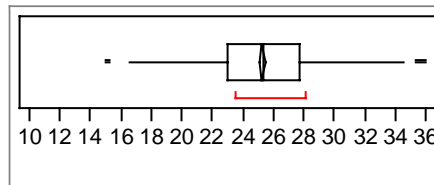
1000 randomizations for improvement with drug A,
Completely randomized design

Mean: 21.3
St. Dev: 3.2



1000 randomizations for improvement with drug B,
Completely randomized design

Mean: 25.3
St. Dev: 3.4



[Technical note: the small bracket in the boxplots marks the “densest half” of the data, the smallest interval that contains half of the data, an additional feature provided by the software. For normal distributions, this corresponds to the interquartile range.]

II. Randomized Complete Block – Blocked by age

Students need to use the table of subjects sorted by age. If you have entered all of the data in the calculator, you do not need to photocopy this table.

Here are boxplots that illustrate the pattern you should see. These pictures were generated using the software JMP-IN 4 by randomizing the subjects 1000 times. The boxplots for your class will be somewhat different since you have a much smaller number of randomizations.

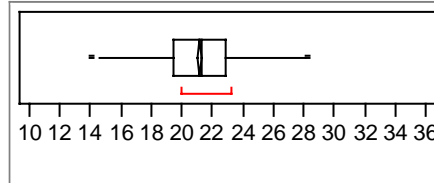
For drug B, the distribution is nearly identical to the completely randomized design. This indicates that blocking by age was not useful in reducing the variability of the sample

means. However, blocking did seem to slightly reduce the variability for drug A, which may come as a surprise. The reasons for this are explained in part VI.

There is still significant overlap, however, which prevents us from being able to make a clear decision in favor of drug B.

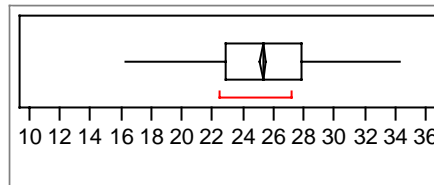
1000 randomizations for improvement with drug A,
Randomized block design (Age)

Mean: 21.2
St. Dev: 2.6



1000 randomizations for improvement with drug B,
Randomized block design (Age)

Mean: 25.3
St. Dev: 3.4

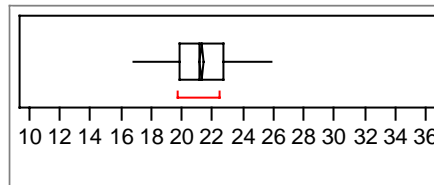


III. Randomized Complete Block – Blocked by initial cholesterol

Here are some typical boxplots, again using 1000 randomizations. The variability in sample means has been reduced due to the blocking on initial cholesterol. In this trial, boxplots of sample means for drug A and drug B should still show a difference in means, but with the reduction in variability there should be much less overlap (less than 5% of the time drug A should be more effective than drug B). This is more convincing evidence that drug B is better than drug A.

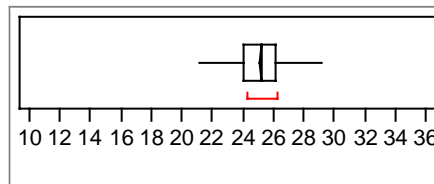
1000 randomizations for improvement with drug A,
Randomized block design (Initial cholesterol)

Mean: 21.3
St. Dev: 1.8



1000 randomizations for improvement with drug B,
Completely randomized design (Initial cholesterol)

Mean: 25.2
St. Dev: 1.5

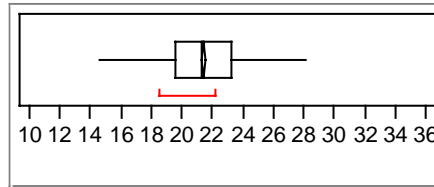


IV. Randomized Complete Block – Blocked by gender and initial cholesterol

Here are some typical boxplots, again using 1000 randomizations. The variability for both drug A and drug B increased slightly, compared to blocking by initial cholesterol only. Blocking by a second variable will often increase the variability if that variable is unrelated to the response. Blocking by gender actually makes it harder to decide that drug B is better so we will ignore it from now on.

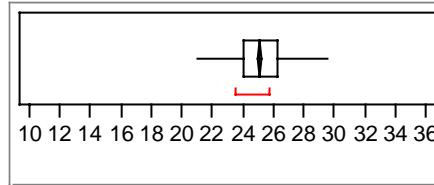
1000 randomizations for improvement with drug A,
Randomized block design (Gender and cholesterol)

Mean: 21.4
St. Dev: 2.5



1000 randomizations for improvement with drug B,
Randomized block design (Gender and cholesterol)

Mean: 25.1
St. Dev: 1.6

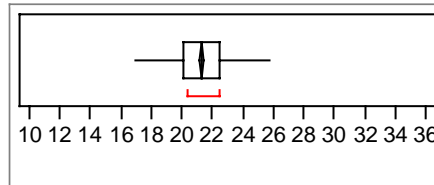


V. Randomized Complete Block – Blocked by exercise level and initial cholesterol

Here are some typical boxplots, again using 1000 randomizations. The variability for this blocking design is reduced even more than by blocking on initial cholesterol alone. This indicates that exercise level is related to the response and including it as a blocking factor makes it obvious that drug B is better than drug A. For this particular data set, no other blocking design will provide a more conclusive result.

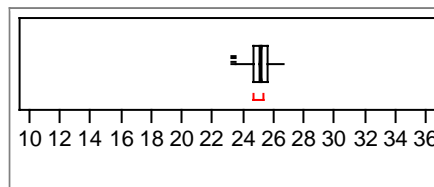
1000 randomizations for improvement with drug A,
Randomized block design (Exercise level and initial cholesterol)

Mean: 21.3
St. Dev: 1.5



1000 randomizations for improvement with drug B,
Randomized block design (Exercise level and initial cholesterol)

Mean: 25.2
St. Dev: 0.6

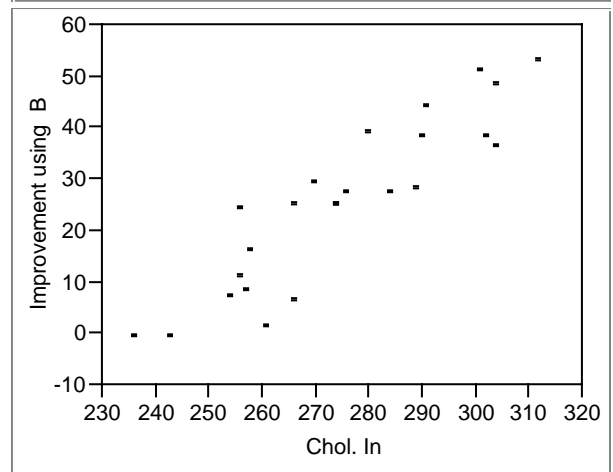
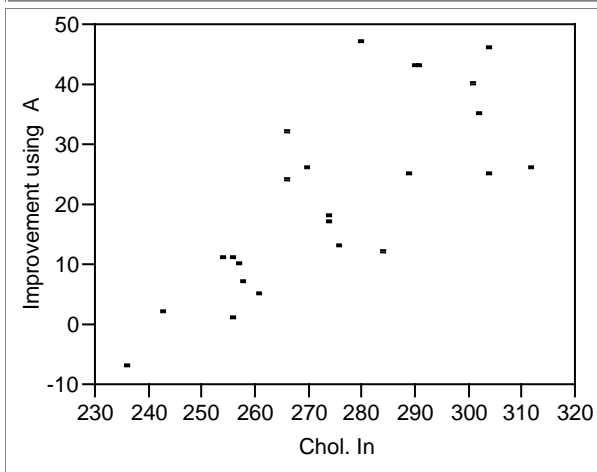
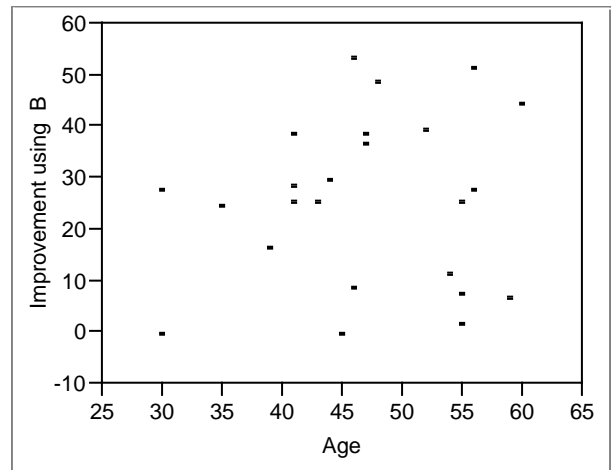
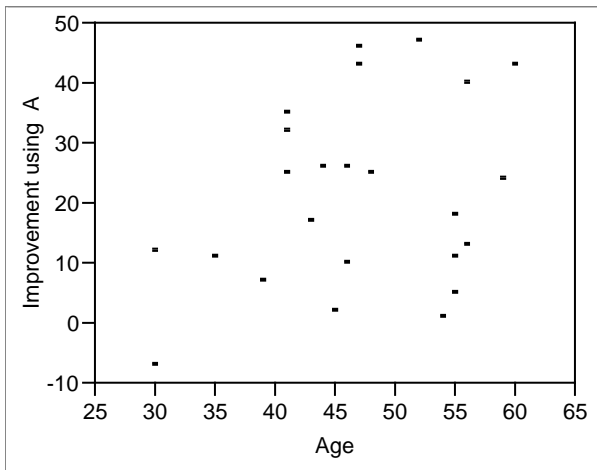


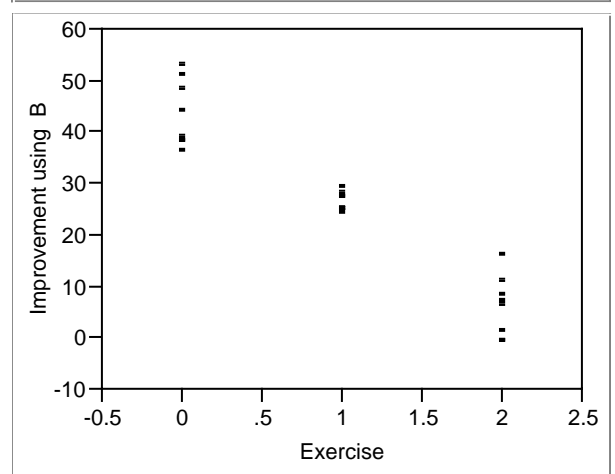
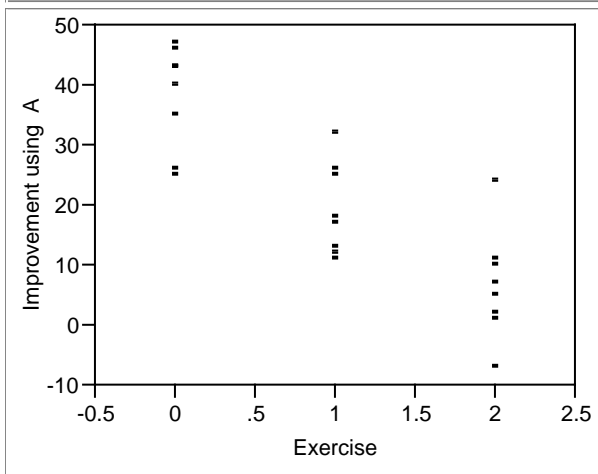
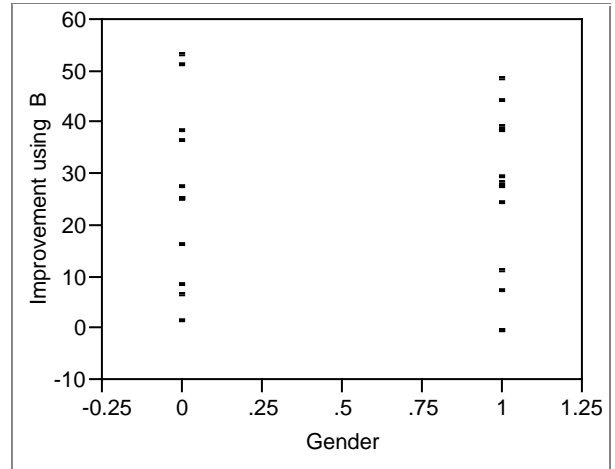
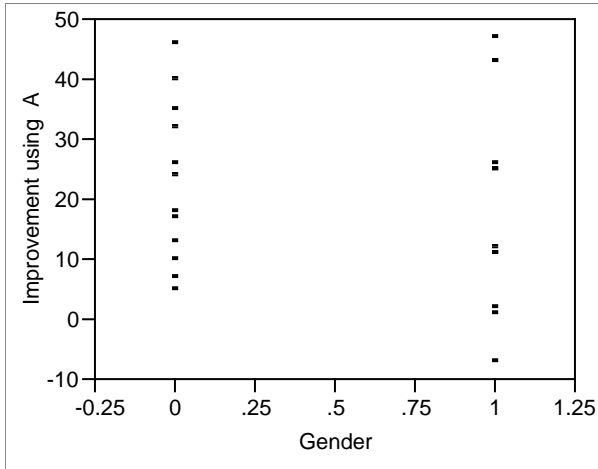
VI. When will blocking be useful?

The scatterplots that the students produce in this section are shown below. Students should observe that the factors that are effective in blocking are those that have a strong association with the response variable.

A few interesting things to note:

- For age, there is a weak association with the improvement using drug A. This is why blocking reduces the variation somewhat for A. This is just a chance occurrence, since the values for age have been randomly selected.
- In the initial cholesterol plots, there is a stronger association with the improvement in drug B than in drug A. This is why the reduction in variation for drug B is greater than the reduction with drug A when blocking on initial cholesterol.
- You may have observed that the variability actually increased when we included gender as a blocking factor together with initial cholesterol. Since gender isn't associated with the response, it shouldn't reduce the variability. In fact, sorting by gender made the initial cholesterol levels in the blocks less similar, therefore increasing the variability from when blocked by initial cholesterol only.
- Finally, since exercise level is strongly associated with the response, adding it as a blocking factor should reduce the variability even further (and it did). Since the association was much stronger with drug B, the variation decreases even more than with drug A.





Are there times when you shouldn't block? Yes, when the potential blocking variable is not associated with the response variable. It's extra work to block, and there is no reward in the form of reduced variability if there is no association.

4. Materials

Materials available at the website include:

1. These teacher notes
2. Student packet
3. Calculator commands for the TI-83
4. The data in a variety of formats
 - Blocking Data (excel)
 - Blocking Data (minitab)
 - Blocking Data (JMP)

5. A program for the TI-83 (Blocking Data) which creates the data table as lists in the calculator.
6. Scripts for JMP-IN 4 to simulate these designs.