Journal of the American Pomological Society 65(1): 2-16 2011

Design of Field Experiments: Influence of Treatment Response Relative to Standard Deviation and Blocking Factor Characteristics on Efficient Blocking Strategy

Ed Stover[1] and Kenneth Portier[2]

Abstract

Selection of experimental design can markedly influence efficiency of field research. This study used Monte Carlo simulations to compare the ability of different field experimental designs to distinguish defined treatment differences, and the paper concludes with a section on practical use of the information obtained. In each simulation, a single experimental treatment was compared to a control treatment and each treatment was applied to twelve trees representing similar research effort. Experimental designs compared were twelve blocks with a single tree per treatment, six blocks of two trees per treatment, four blocks of three trees per treatment, and two blocks with six trees per treatment. In each case, analyses were compared in which data were collected with single tree experimental units (multiple trees independently assigned the same treatment within each block) or as is often done with spatial blocking, data were pooled on a group of trees (one data point per multiple tree experimental unit). Trees were blocked according to a specified factor, which was quantified for these comparisons but could represent a qualitative factor such as spatial position. The probability of rejecting the null hypothesis was computed for a range of situations including small and large values for the following parameters: treatment response, standard deviations of the response, blocking factor effects, and blocking factor standard deviations. In all cases, the probability of rejecting a false null hypothesis was significantly greater when data were collected on single-tree experimental units, and decreased as the number of trees pooled per data point increased (and number of blocks decreased). When data were collected on single-tree experimental units and the factor used for blocking actually had no relationship to the response variable, all four designs had similar probabilities of rejecting the null hypothesis; however, power decreased with increasing block size (more trees per block but fewer blocks) when the blocking factor was significantly correlated with the response variable but treatment did not change the slope of the blocking factor vs. response variable. When the blocking factor effect was significant and there was a significant treatment-by-block interaction, use of a single tree per treatment per block had the least power, but power decreased substantially with block size greater than two trees per treatment. In the last case, failing to account for block by treatment interaction effects resulted in test statistics having little power to reject the null hypothesis even when treatment effects were strong. These analyses indicate that use of one or two trees per treatment per block with data collected on individual-tree experimental units provides the greatest efficiency in distinguishing treatment effects, and that two trees per treatment per block is superior when there is a significant treatment-by-blocking factor interaction. When the blocking factor displayed a distinct spatial trend, incorrectly using individual tree data from multi-tree experimental units as pseudo-replicates resulted in false rejections of the null hypothesis well beyond the specified α=0.05, sometimes approaching P=0.50. Researchers are cautioned that proper analysis of multi-tree experimental units yields the same F-test using individual subsample data or single mean values representing each collective experimental unit.

Many issues of current importance to production agriculture can only be addressed through field experiments, in which the effects of various treatments are compared under controlled conditions. Since field-grown plants are often influenced by site variables or other factors that can be predicted prior to experimentation, the randomized complete block design (RCBD) is one of the most frequently used experimental designs for agricultural field studies (2). Researchers vary widely in their approach to design of such experiments, with many devoting extensive effort to identification and quantification of potential blocking factors, while others may simply assume that field position will be the dominant influence.

[1] USDA, ARS US Horticultural Research Laboratory, 2001 S. Rock Rd., Ft. Pierce, FL 34945
[2] University of Florida, Dept. of Statistics, Institute of Food and Agricultural Sciences, PO Box 110339, Gainesville, FL 32611-0339

Because of the high degree of variability seen in field-grown tree species, some researchers have adopted the strategy of pooling data on multiple tree experimental units to buffer variability, and using fewer experimental units. However, if blocking is not really needed, because the factors used to create blocks are not related to the response of interest, then the effort and degrees of freedom expended on the blocking factor can reduce the power of the test for treatment effects.

It is often not clear how blocking affects our ability to successfully distinguish treatments that produce modest differences from the non-treated controls. The objective of this research was to quantify the impacts of blocking strategy, statistical characteristics of the blocking factor, and standard deviation of the response variable, on the power of a simple two-treatment comparison. In this way, we hope to emphasize and understand the importance of a carefully considered blocking strategy. Furthermore, in this era in which cutbacks have diminished statistical support for researchers and availability of easily used software permits most research to be conducted without statistical consultation, it is hoped that this information will prove especially useful to field horticulturists who independently design and analyze their experiments. The end section of the paper, "*Using this Information",* is provided as a summary for those interested in field experimentation but less interested in the statistical justification for the recommendations that are made as a result of this study,

Materials and Methods

Variability inherent in field experiments results in experimentwise variability in response to a given treatment, sometimes including incorrect rejection of the null hypothesis, even when the experiment is conducted under seemingly identical conditions. For this reason, meaningful comparison of experimental designs requires numerous repetitions of the same experiment. The number of repeated trials needed for confident comparison of designs is so great that actual field comparisons may be impractical. Fortunately, computer simulations permit numerous experimental runs under defined conditions, providing valuable inferences on responses likely to be observed in the field.

Monte Carlo simulations were performed using SAS® (SAS Institute Cary, N.C.). A normally distributed blocking factor was incorporated into the general linear model used to generate the data. Blocking factor values were used as the basis for blocking of the units prior to analysis. To run these simulations, it was necessary to quantify the relationship between the blocking factor and the response variable; however, this blocking factor is also intended to represent more qualitative approaches such as blocking by spatial orientation in the field. In addition, a treatment-by-blocking-factor interaction term was modeled and its effect on the power of the test for treatment effects was measured. A Type I error rate of $P<0.05$ was used for all analyses, except where noted.

Models. The general linear model used to generate the responses for the control group units is defined by the following relationship.

$$Y_i = \alpha_0 + \alpha_1 X_i + \varepsilon_i$$

The blocking factor, X, is defined for all simulations as having a normal distribution with mean zero and standard deviation, s_X. Then α_0 represents the overall mean response at the mean level of the blocking factor X. The coefficient α_1 measures the strength of the relationship between the response Y and the blocking factor X. The residual term ε is assumed to have the same normal distribution for both control and treatment groups with common standard deviation s_ε. For the treated units, the linear model is given as follows:

$$Y_i = (\alpha_0 + \gamma_0) + (\alpha_1 + \gamma_1) X_i + \varepsilon_i$$

In addition to the components described for the control model, γ_0 measures the overall effect of the treatment on the mean response

at the average level of the blocking factor and γ_1 measures the change in response to the blocking factor produced by the addition of the treatment. γ_0 is set as a percentage of α_0.

Factors common to all simulations. For all experiments, the value of α_0, mean value of the response variable in the control treatment, was set at 100 to permit easy assessment of incremental improvement from experimental treatments (γ_0). Each experiment was run as a factorial with three values each for s_X (5, 10, 20), and s_ε (5, 10, 20). When the common standard deviation $s_\varepsilon = 20$, the value of the response variable will range from about 60 to 140 among many individual control trees assessed, while the range for $s_\varepsilon = 5$ would only be 90 to 110. Similarly, variability in the response variable due to the blocking factor would be 60 to 140 for $s_X = 20$, but only 90 to 110 for $s_X = 5$.

For each of the scenarios, 5000 simulated experimental runs were generated with comparisons between seven different blocking designs (Table 1). When each treatment is applied to more than one tree in each block we compare the use of multi-tree experimental units, as in a standard RCBD, to multiple single-tree experimental units which are independently randomized within the block, a design known as a Generalized RCBD (Fig. 1). In all cases the total number of experimental units is 24 (12 per treatment), permitting comparison of blocking strategies with the same amount of effort expended on data collection. While 24 trees represents a very small experiment, use of 12 trees per treatment is relatively large for the field experiments we are interested

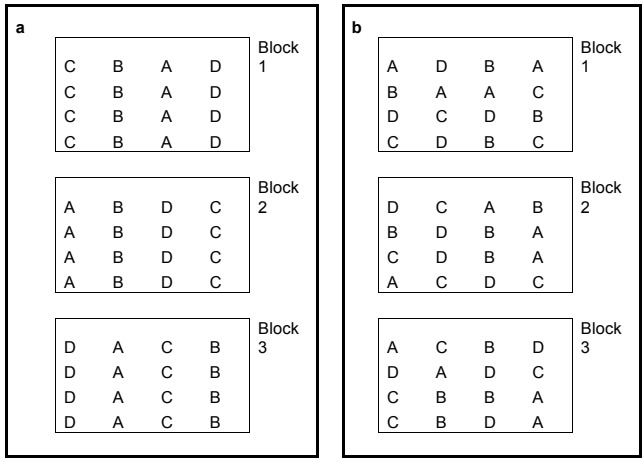

Fig. 1. Depiction of two approaches to a Randomized-Complete-Block (RCBD) design with multiple trees receiving each of four treatments (A, B, C, and D) in each of three blocks. Fig. 1a illustrates the standard RCBD, in which a single four-tree experimental unit receives each treatment. The proper error term for the F=test is Block x Treatment with 6 degrees of freedom [(# Blocks-1) x (# Treatments-1)]. If each of the trees assigned to a treatment as a group is analyzed as an independent experimental unit, then this is pseudoreplication. Fig. 1b illustrates the generalized RCBD, in which four single-tree experimental units are independently randomized within each block for each treatment. The proper error term for the F-test is the default MSE, which is Tree (Block x Treatment) which has 36 degrees of freedom [(# Blocks) x (# Treatments) x (# Trees per treatment per block -1)].

Table 1. Factors common to all simulations.

Treatments	Number of simulated experimental iterations	Values for α_0, the mean for response variable in controls	Values for s_e, the standard error of response variable	Values for s_x, the standard error of blocking factor	Number of blocks	Trees / treatment / block	Error term for F-test with collection of data from individual trees	Error term for F-test with pooled collection of data from groups of trees
control and treated	5000	100	5, 10, & 20	5, 10, & 20	12	1	block x tree MS	not applicable
					6	2	MSE[z]	block x tree MS[y]
					4	3		
					2	6		

[z] Appropriate for analyzing data from a generalized RCBD in which multiple trees are independently assigned the same treatment within each block
[y] Appropriate for analyzing data from an RCBD in which multiple trees in a group are assigned the same treatment within each block

in simulating, which normally involve comparison of more treatments, but it simplifies discussion to compare a single treatment to a non-treated control. Within the constraints defined for each experiment, setting values for different model parameters allows us to test different hypotheses.

Except for the simulation in Experiment 5 (Table 8), the variability within each block is random and adjacent trees are no more likely to be similar than are any two trees within an individual block.

Simulations in which treatment has no effect. For all runs of the simulation in Experiment 1, the value of γ_0 was set at 0, indicating that there is no effect of treatment on the response variable and γ_1 was also 0, indicating that the experimental treatment does not affect the slope of the response variable *vis-à-vis* the blocking factor. The slope of the response variable vs. the blocking factor, α_1, was tested at 0, 0.5, and 1.0 (see Fig. 2). Results are discussed but not shown.

Simulations in which treatment affects mean but blocking factor has no effect. The effect of treatment on the response variable, γ_0, was compared at 5 and 10% in Experiment 2. The value of α_1 and γ_1 were set at 0, indicating no relationship between the blocking factor and the response variable (see Fig. 3; Table 3).

Simulations in which treatment affects mean but not slope of blocking factor vs. response variable, and blocking factor affects response variable. The effect of treatment on the response variable, γ_0, was compared at 5% and 10% in Experiment 3. The value of α_1, which is the slope between the response variable and the blocking factor, was set at 0.5 and 1.0 indicating moderate and large influence of the blocking factor on the response variable. γ_1 was set at 0, indicating no effect of treatment on the slope between the blocking factor and the response variable (see Fig. 4; Tables 4 and 5).

Simulations in which treatment affects mean and slope of blocking factor vs. response variable. The effect of treatment on the response variable, γ_0, was set at 5% and α_1, which is the slope between the response variable and the blocking factor, was set at 0.5 in Experiment 4. All comparisons were conducted with γ_1,

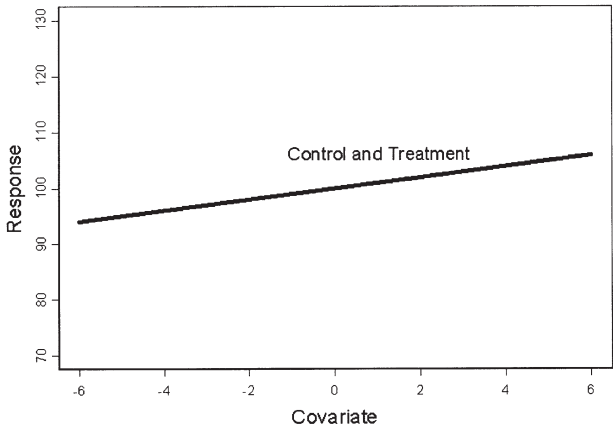

Fig. 2. No treatment effect but with blocking factor association (α_0 = 100, γ_0 = 0, α_1=1.0, γ_1=1.0).

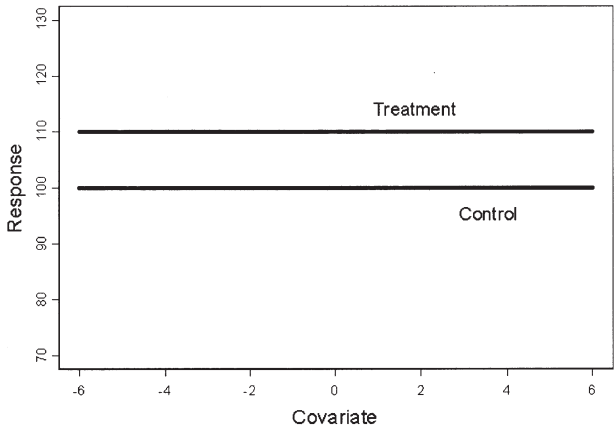

Fig. 3. Treatment affects mean but blocking factor has no effect (α_0 = 100, γ_0 = 10, α_1=0.0, γ_1=0.0).

Table 2. Factors tested in each experiment.

Experiment	Values for γ_0, the effect of treatment on the response variable	Values for α_1, the slope of the response variable vs. the blocking factor	Values for γ_1, the effect of treatment on the slope of the response variable vs. the blocking factor	Relationship between value of blocking factor covariate (α_1) and assignment of experimental units	Table(s) listing results
1- no effect of treatment on response variable (see Fig. 2)	0	0, 0.5, 1.0	0	Each block is uniformly variable and adjacent trees are likely no more alike than any two trees in the same block	data not shown
2- no blocking factor effect but treatment increases response variable by 5 and 10% (see Fig. 3)	5, 10	0	0	As above	Table 3
3- blocking factor vs. response variable has moderate to large slope, treatment increases response variable by 5 and 10% but doesn't influence slope of response variable vs. blocking factor (see Fig. 4)	5, 10	0.5, 1.0	0	As above	Tables 4 & 5
4- blocking factor vs. response variable has moderate slope, treatment increases response variable by 5 % but also influences slope of response variable vs. blocking factor (see Fig. 5)	5	0.5	-0.5, -0.1, 0.1, 0.5	As above	Tables 6 & 7
5- blocking factor vs. response variable has no, moderate, or large slope, treatment increases response variable by 0, 5 or 10% but doesn't influence slope of response variable vs. blocking factor (see Fig. 4)	0, 5, 10	0, 0.5, 1.0	0	Effect of blocking factor generally increases from one end of the block to the other, so that adjoining trees are more similar to each other.	Table 8

effect of treatment on the slope between the blocking factor and the response variable, at –0.1, 0.1, -0.5, and 0.5, to determine how the power of the test under different blocking strategies is affected when there is a significant treatment-by-block interaction (see Fig. 5; Tables 6 and 7).

Simulations in which the slope of blocking factor vs. response variable is significant and the blocking factor displays a spatial relationship. The blocking factor was arranged in an escalating linear fashion and assigned to treatments before being affected by the standard deviation for the blocking factor in Experiment 5. As a result, in this experiment, adjoining trees were usually more similar to each other than to trees at the far end of the block, which is likely to be representive of many field experiments with spatial blocking.

Most comparisons were made with no treatment effect on slope between blocking factor and response variable (γ_1=0; essentially there is no treatment-by-block interaction; demonstrated in Fig. 4). No experimental increment (treatment has no effect on the response variable) was compared with no or modest slope of block vs. response variable (α_1=0, 0.5). A small experimental increment (treatment increases response variable by 5%) was compared with no, modest, or large slope of block vs. response variable (α_1=0, 0.5, 1.0) (Table 8).

Using a blocking factor simulating a spatial relationship, a number of comparisons were conducted in which treatment does affect the slope between the blocking factor and the response variable (γ_1 at –0.1, 0.1, -0.5, and 0.5; Fig. 5). These data are briefly discussed but not shown.

Testing. In the simulations, two residual error terms were used in the testing where there were multiple trees per treatment per block. In the first scenario, each tree is used as an experimental unit, as would be properly done in a generalized RCBD when each treatment is individually assigned to a separate tree, even if there are multiple trees of the same treatment in each block, and the residual error is properly estimated by tree-to-tree variation pooled across each block by treatment. Therefore, the treatment mean squares (MS) is divided by the error MS from the ANOVA table to generate

the appropriate F test. In the second scenario, treatments are assigned randomly to groups of trees within each block (e.g. a spatial block of 1, 2, 3, or 6 trees) and data are pooled across all trees in each experimental unit (e.g. data collected on yield of all fruit from a three tree experimental unit) in which case the proper F test is generated by dividing the treatment MS by the block x treatment MS from the ANOVA. SAS code for these two approaches can be found in Schabenberger (6). When there is a single tree per treatment per block the error term for the F test can only be generated by dividing the treatment MS by the block x treatment MS from the ANOVA.

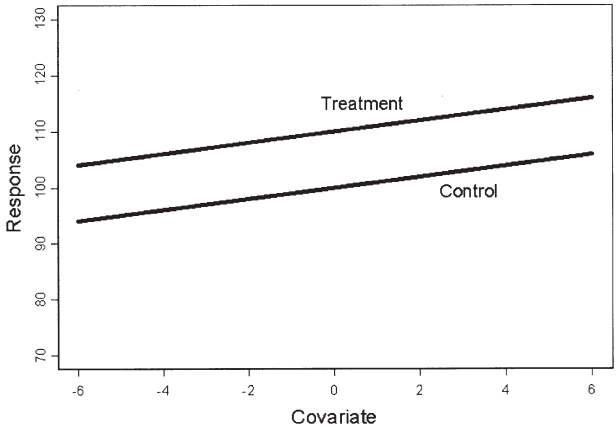

Fig. 4. Treatment affects mean and there is a blocking factor effect, but the blocking factor effect is not affected by treatment (α_0 = 100, γ_0 = 10, α_1=1.0, γ_1=1.0).

Results and Discussion

The Monte Carlo approach provides an outcome for each simulated experiment indicating the proportion of treatment tests in which the null hypothesis is rejected. Power is defined as the probability of correctly rejecting the null hypothesis, and is therefore the probability of distinguishing between treatment effects that would be seen with a sufficiently large and properly planned experiment. When a single tree per treatment is used in each block, the residual variability can only be estimated using the treatment-by-block interaction. For smaller numbers of blocks (those with multiple trees per treatment per block in these simulations), both treatment-by-block and residual variability estimates are available, but are appropriate for different experimental designs.

Since proper experimental design demands use of sufficient replications to make a finding of significance likely, given expected treatment effects and variability (3), some of the discussion focuses on only those results in which at least one of the experimental designs provided power greater than 20%.

Predicted responses support validity of method. The output from this analysis included many responses that one would expect for a properly functioning model system. When both γ_0 (effect of treatment on response variable) and γ_1 (effect of treatment on slope of response variable vs. blocking factor) are set at zero, *P*r (Reject) averaged 0.050 which was the type I error rate (probability of incorrectly rejecting the null hypothesis) designated for this analysis. In individual groups of 5000 simulations, the type I error rate ranged from 0.042 – 0.059, with no apparent differences in error rates among different blocking strategies or other experimental parameters, when the variability within each block is random and adjacent trees are no more likely to be similar than are any two trees within an individual block (all experiments except Experiment 5; data not shown).

Also as one would expect, the probability of correctly rejecting the null hypothesis was greater as the effect of treatment on response variable increased relative to response variable standard error (Tables 3 to 7).

Simulations in which treatment affects the response variable but blocking factor has no effect. There is no benefit in blocking when the blocking factor has no effect on the response variable, but if this isn't known prior

Table 3. Simulation results with no blocking factor effect and experimental increment of 5 and 10 percent.

					12 blocks of 1 tree / treatment	6 blocks of 2 trees / treatment		4 blocks of 3 trees / treatment		2 blocks of 6 trees / treatment	
					P of rejecting null hypothesis						
						Group of trees pooled for data	Individual tree data	Group of trees pooled for data	Individual tree data	Group of trees pooled for data	Individual tree data
Trial	Slope of response variable vs. blocking factor (α_1)	Treatment effect on response variable (γ_0)	Standard error of response variable (s_e)	Standard error of blocking factor (s_x)	Block × trt as error for F test	Block × trt as error for F test[z]	MSE as error for F test[y]	Block × trt as error for F test[z]	MSE as error for F test[y]	Block × trt as error for F test[z]	MSE as error for F test[y]
1A	0	5	5	5	0.62	0.51	0.62	0.39	0.62	0.14	0.64
2A				10	0.61	0.50	0.61	0.39	0.63	0.15	0.64
3A				20	0.61	0.49	0.60	0.40	0.63	0.16	0.65
4A			10	5	0.20	0.17	0.21	0.14	0.21	0.09	0.22
5A				10	0.21	0.18	0.21	0.14	0.21	0.08	0.21
6A				20	0.20	0.17	0.21	0.14	0.21	0.09	0.21
7A			20	5	0.08	0.08	0.09	0.07	0.08	0.06	0.09
8A				10	0.10	0.08	0.08	0.07	0.09	0.06	0.09
9A				20	0.09	0.08	0.09	0.07	0.09	0.06	0.09
10A		10	5	5	0.99	0.97	0.99	0.89	1.00	0.30	1.00
11A				10	0.99	0.97	0.99	0.89	1.00	0.30	1.00
12A				20	0.99	0.97	0.99	0.89	1.00	0.30	1.00
13A			10	5	0.62	0.51	0.61	0.40	0.64	0.15	0.63
14A				10	0.60	0.50	0.61	0.39	0.64	0.14	0.65
15A				20	0.62	0.52	0.62	0.39	0.63	0.16	0.64
16A			20	5	0.19	0.17	0.20	0.14	0.21	0.08	0.22
17A				10	0.20	0.17	0.20	0.14	0.21	0.08	0.21
18A				20	0.20	0.18	0.21	0.15	0.21	0.08	0.22

[z] Appropriate for analyzing data from an RCBD in which multiple trees in a group are assigned the same treatment within each block

[y] Appropriate for analyzing data from a generalized RCBD in which multiple trees are independently assigned the same treatment within each block

to conducting the experiment, what is the effect of having selected different blocking strategies? In the generalized RCBD, where several single-tree replicates are individually randomized within each block, the appropriate estimate for error variance in the F test is the mean square error term, commonly known as sampling error (which is adjusted for block effects) and there was no effect of blocking method on experimental power. However, if multiple-tree experimental units are used and there is no true replication of a treatment within a block, the block x treatment interaction mean square is used in the F test, and the power decreases as number of blocks decreases (Table 3).

Simulations in which treatment affects the response variable but not the slope of the blocking factor vs. response variable. When the blocking factor significantly influences the response variable and the standard error of the blocking factor is equal to or greater than the residual standard error, there is often a considerable benefit to using more blocks. In this scenario, with no significant treatment-by-block interaction (Fig. 4), using more and smaller blocks results in higher power than fewer and larger blocks. This is true whether the association is moderate (Table 4) or strong (Table 5). In our simulations, as the slope of

Table 4. Simulation results with moderate blocking factor effect and experimental increment of 5 and 10 percent. There is no treatment effect on slope between blocking factor and response variable (γ_1=0).

					12 blocks of 1 tree / treatment	6 blocks of 2 trees / treatment		4 blocks of 3 trees / treatment		2 blocks of 6 trees / treatment	
					P of rejecting null hypothesis						
						Group of trees pooled for data	Individual tree data	Group of trees pooled for data	Individual tree data	Group of trees pooled for data	Individual tree data
Trial	Slope of response variable vs. blocking factor (α_1)	Treatment effect on response variable (γ_0)	Standard error of response variable (s_e)	Standard error of blocking factor (s_x)	Block × trt as error for F test	Block × trt as error for F test[z]	MSE as error for F test[y]	Block × trt as error for F test[z]	MSE as error for F test[y]	Block × trt as error for F test[z]	MSE as error for F test[y]
1B	0.5	5	5	5	0.58	0.46	0.56	0.33	0.55	0.12	0.42
2B				10	0.53	0.36	0.44	0.24	0.36	0.08	0.2
3B				20	0.38	0.21	0.23	0.12	0.17	0.05	0.09
4B			10	5	0.20	0.16	0.2	0.13	0.20	0.08	0.18
5B				10	0.19	0.16	0.19	0.12	0.17	0.07	0.14
6B				20	0.17	0.13	0.15	0.10	0.12	0.06	0.09
7B			20	5	0.08	0.08	0.08	0.08	0.09	0.05	0.08
8B				10	0.09	0.08	0.09	0.07	0.08	0.05	0.08
9B				20	0.08	0.07	0.08	0.07	0.08	0.05	0.07
10B		10	5	5	0.99	0.95	0.99	0.83	0.99	0.22	0.93
11B				10	0.98	0.88	0.95	0.63	0.90	0.15	0.61
12B				20	0.90	0.61	0.70	0.33	0.53	0.08	0.22
13B			10	5	0.61	0.50	0.61	0.36	0.61	0.14	0.56
14B				10	0.59	0.47	0.56	0.33	0.53	0.12	0.39
15B				20	0.53	0.35	0.43	0.24	0.37	0.08	0.19
16B			20	5	0.21	0.17	0.21	0.14	0.20	0.07	0.21
17B				10	0.21	0.16	0.19	0.14	0.20	0.07	0.18
18B				20	0.19	0.15	0.19	0.13	0.18	0.07	0.14

[z] Appropriate for analyzing data from an RCBD in which multiple trees in a group are assigned the same treatment within each block

[y] Appropriate for analyzing data from a generalized RCBD in which multiple trees are independently assigned the same treatment within each block

the blocking factor effect on response increases, more noise is introduced into the data and hence the power of most tests decreases. With a stronger blocking factor effect, loss of power is much greater as the number of blocks decreases.

Use of multi-tree experimental units, which demands use of the treatment-by-block interaction term as the residual error estimate in the F test, never produces a more powerful test than taking data from individual-tree experimental units (individually randomized within the block) where the appropriate error term is the residual mean square error term (sampling error).

Simulations in which treatment affects the response variable and the slope of the blocking factor vs. response variable. When the treatment influences the slope of the blocking factor vs. the response variable (meaning there is an interaction between the block and the treatment; Fig. 5), effect of experimental design was strongly influenced by degree of slope alteration. Small treatment influences on slope (Table 6) resulted in similar power for the designs with 12 blocks of one tree per treatment and 6 blocks of two trees per treatment. However, when treatment substantially altered

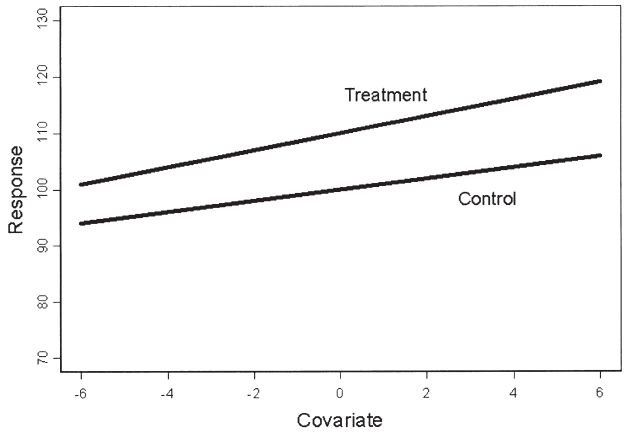

Fig. 5. Treatment affects mean slope of blocking factor (α_0 = 100, γ_0 = 10, α_1=1.0, γ_1=0.5).

the slope of response variable vs. blocking factor, there was a major advantage of collecting data on multiple single-tree experimental units for each treatment in a block (Table 7). This makes sense because data on more than one tree per treatment per block makes it possible to separate treatment-by-block interaction and remove this source of variability from the error term (6). Note that there was never an advantage in having more than two trees per treatment per block (Table 7). For maximum efficiency, it appears that only the minimal amount of within block replication needed to estimate a true residual error should be used.

The use of pooled multiple tree data collection again reduced power, but the relationship among experimental designs was quite different. With a small effect of treatment on slope of response variable vs. blocking factor (γ_1=0.1 or –0.1, Table 6), significant but modest loss of power occurred when use of multiple-tree experimental units was contrasted with use of single-tree units for each design. However, for greater treatment effect on slope (γ_1=0.5 or –0.5, Table 7), all designs had extremely low probabilities of rejecting the null hypothesis when pooled multiple-tree data were used.

Simulations in which the slope of blocking factor vs. response variable is significant and the blocking factor displays a spatial relationship. In many field experiments with spatial blocking, adjoining trees may sometimes be much more alike than two more distant trees within the same block. When this was simulated by arranging the blocking factor in an escalating fashion for treatment assignment, which is only appropriate for the standard RCBD, the results were quite interesting. In this experiment (number 5, table 2), Pr (Reject) appropriately hovered around 0.05 when both γ_0 (effect of treatment on response variable) and α_1(slope of response variable vs. blocking factor uninfluenced by treatment) are set at zero (Table 8, Lines 1F through 6F), but when γ_0=0 and and α_1 was positive (there was a blocking effect) then the Pr (Reject) ranged from 0.02 to 0.12, when the appropriate error term (block x treatment, Table 8) was used. Of greater concern is that with no treatment effect, significant block effect on response variable, and use of the inappropriate error term (MSE, used as though each tree is an independent experimental unit even though assigned as a group), the null hypothesis is incorrectly rejected in a rather high proportion of trials (see lines 10F through 15 F in Table 8).

Use of individual tree data from trees which were not independently assigned to treatments is known as pseudoreplication (1). This violates proper statistical practice, and can result in a high probability of inaccurate statistical inference. This would occur when similarity between adjoining plants results in a marked underestimate of tree to tree variability. When pseudoreplicates are used where a blocking effect varies across the block as well as between blocks, and there is a treatment effect, use of multi-tree experimental units does increase the likelihood of finding the treatments significantly different (compare MSE as error term in lines 21F with no block

Table 5. Simulation results with large blocking factor effect and experimental increment of 5 and 10 percent. There is no treatment effect on slope between blocking factor and response variable (γ_1=0).

					12 blocks of 1 tree / treatment	6 blocks of 2 trees / 4 blocks of 3 trees / 2 blocks of 6 trees / treatment					
					P of rejecting null hypothesis						
						Group of trees pooled for data	Individual tree data	Group of trees pooled for data	Individual tree data	Group of trees pooled for data	Individual tree data
Trial	Slope of response variable vs. blocking factor (α_1)	Treatment effect on response variable (γ_0)	Standard error of response variable (s_e)	Standard error of blocking factor (s_x)	Block × trt as error for F test	Block × trt as error for F test[z]	MSE as error for F test[y]	Block × trt as error for F test[z]	MSE as error for F test[y]	Block × trt as error for F test[z]	MSE as error for F test[y]
1C	1	5	5	5	0.53	0.36	0.44	0.24	0.36	0.08	0.20
2C				10	0.37	0.20	0.24	0.12	0.17	0.06	0.09
3C				20	0.18	0.10	0.12	0.07	0.09	0.05	0.06
4C			10	5	0.19	0.16	0.18	0.12	0.18	0.06	0.14
5C				10	0.18	0.13	0.14	0.10	0.14	0.06	0.08
6C				20	0.13	0.08	0.09	0.06	0.08	0.05	0.06
7C			20	5	0.08	0.07	0.09	0.07	0.08	0.06	0.08
8C				10	0.09	0.08	0.08	0.06	0.08	0.05	0.07
9C				20	0.08	0.07	0.08	0.05	0.07	0.05	0.06
10C		10	5	5	0.98	0.87	0.95	0.64	0.90	0.39	0.59
11C				10	0.90	0.59	0.70	0.34	0.51	0.08	0.22
12C				20	0.56	0.27	0.29	0.14	0.19	0.06	0.09
13C			10	5	0.58	0.46	0.56	0.33	0.54	0.11	0.41
14C				10	0.52	0.36	0.44	0.22	0.36	0.07	0.19
15C				20	0.37	0.20	0.24	0.13	0.18	0.06	0.09
16C			20	5	0.20	0.17	0.20	0.13	0.20	0.07	0.19
17C				10	0.19	0.15	0.18	0.13	0.17	0.07	0.13
18C				20	0.17	0.13	0.15	0.09	0.13	0.06	0.08

[z] Appropriate for analyzing data from an RCBD in which multiple trees in a group are assigned the same treatment within each block

[y] Appropriate for analyzing data from a generalized RCBD in which multiple trees are independently assigned the same treatment within each block

effect to 27F with a strong block effect in Table 8), but this is largely a spurious effect from using the wrong error term, and largely reflects the probability of erroneously finding even a non-significant treatment effect to be significant (again see lines 10F through 15 F in Table 8). The danger of accidentally analyzing samples as pseudoreplicates can be avoided by conducting the analysis on the means for multi-tree experimental units, since there should be no difference in the F-test for use of means or individual samples in a properly conducted test. When the correct error term (Block x Treatment) is used for an RCBD without multiple independent experimental units/treatment/block , the highest probability of rejecting a false null hypothesis occurs from using more blocks. In this study this is achieved using 12 blocks each with a single tree per treatment per block (compare Block x Treatment as error term in lines 16F through 33F in Table 8).

When the treatment significantly influences the slope of the response variable *vis-à-vis* the blocking factor (a significant block x treatment interaction), the same spurious effects are observed from use of pseudoreplicates, but the benefits of using two true replicates per

Table 6. Simulation results with medium blocking factor effect and small treatment effect on slope between blocking factor and response variable and experimental increment of 5 percent.

						12 blocks of 1 tree / treatment	6 blocks of 2 trees / 4 blocks of 3 trees / treatment					2 blocks of 6 trees / treatment	
						P of rejecting null hypothesis							
							Group of trees pooled for data	Individual tree data	Group of trees pooled for data	Individual tree data	Group of trees pooled for data	Individual tree data
Trial	Slope of response variable vs. blocking factor (α_1)	Change in slope of response variable vs. blocking factor (γ_1)	Treatment effect on response variable (γ_0)	Standard error of response variable (s_e)	Standard error of blocking factor (s_x)	Block × trt as error for F test	Block × trt as error for F test[z]	MSE as error for F test[y]	Block × trt as error for F test[z]	MSE as error for F test[y]	Block × trt as error for F test[z]	MSE as error for F test[y]
1D	0.5	0.1	5	5	5	0.55	0.41	0.54	0.28	0.51	0.10	0.38
2D					10	0.45	0.26	0.41	0.16	0.34	0.06	0.18
3D					20	0.23	0.11	0.23	0.07	0.17	0.04	0.09
4D				10	5	0.20	0.16	0.20	0.12	0.19	0.07	0.18
5D					10	0.18	0.14	0.18	0.10	0.18	0.06	0.13
6D					20	0.13	0.09	0.15	0.06	0.13	0.05	0.09
7D				20	5	0.09	0.08	0.09	0.07	0.08	0.06	0.08
8D					10	0.08	0.08	0.08	0.06	0.09	0.06	0.08
9D					20	0.08	0.06	0.09	0.06	0.09	0.05	0.07
10D		-0.1		5	5	0.56	0.43	0.56	0.34	0.55	0.10	0.44
11D					10	0.46	0.27	0.47	0.16	0.40	0.07	0.22
12D					20	0.25	0.10	0.28	0.07	0.21	0.05	0.10
13D				10	5	0.19	0.16	0.20	0.13	0.21	0.08	0.19
14D					10	0.18	0.14	0.20	0.10	0.18	0.06	0.14
15D					20	0.14	0.09	0.16	0.07	0.13	0.04	0.10
16D				20	5	0.08	0.08	0.09	0.08	0.09	0.06	0.09
17D					10	0.08	0.07	0.08	0.07	0.09	0.06	0.09
18D					20	0.08	0.06	0.09	0.06	0.08	0.04	0.07

[z] Appropriate for analyzing data from an RCBD in which multiple trees in a group are assigned the same treatment within each block
[y] Appropriate for analyzing data from a generalized RCBD in which multiple trees are independently assigned the same treatment within each block

treatment per block becomes apparent (data not shown) as described earlier for experiment 4 (Table 2) and summarized in Table 7.

Consideration of alternate thresholds for significance and other statistical designs. Where revolutionary increases (20% or greater) from treatments are expected, almost any experimental design will provide sufficient power to distinguish between treatments (including the control). However, most advances

in agricultural science are now incremental and selection of experimental design may substantially influence the power of the test (e.g. comparison 1C in Table 5). Selection of the most efficient experimental design is an obligation for researchers conducting expensive experiments on a limited budget, often with considerable reliance on public funds.

The analyses presented in this study are conducted using the arbitrary threshold of $P<0.05$

Table 7. Simulation results with medium blocking factor effect and medium treatment effect on slope between blocking factor and response variable and experimental increment of 5 percent.

						12 blocks of 1 tree / treatment	6 blocks of 2 trees / treatment		4 blocks of 3 trees / treatment			2 blocks of 6 trees / treatment		
							P of rejecting null hypothesis							
							Group of trees pooled for data	Individual tree data	Group of trees pooled for data	Individual tree data		Group of trees pooled for data	Individual tree data	
Trial	Slope of response variable vs. blocking factor (α_1)	Change in slope of response variable vs. blocking factor (γ_1)	Treatment effect on response variable (γ_0)	Standard error of response variable (s_e)	Standard error of blocking factor (s_b)	Block × trt as error for F test	Block × trt as error for F test[z]	MSE as error for F test[y]	Block × trt as error for F test[z]	MSE as error for F test[y]	Block × trt as error for F test[z]	Block × trt as error for F test[z]	MSE as error for F test[y]	
1E	0.5	0.5	5	5	5	0.23	0.06	0.50	0.02	0.45	0.01	0.28		
2E					10	0.07	0.01	0.37	0.00	0.29	0.00	0.14		
3E					20	0.02	0.01	0.28	0.00	0.20	0.01	0.10		
4E				10	5	0.13	0.07	0.20	0.05	0.20	0.03	0.17		
5E					10	0.07	0.02	0.22	0.01	0.19	0.01	0.12		
6E					20	0.03	0.01	0.22	0.00	0.17	0.01	0.09		
7E				20	5	0.07	0.06	0.09	0.05	0.09	0.04	0.08		
8E					10	0.06	0.04	0.10	0.02	0.10	0.02	0.09		
9E					20	0.03	0.01	0.13	0.01	0.11	0.01	0.08		
10E		-0.5		5	5	0.23	0.06	0.58	0.01	0.57	0.00	0.49		
11E					10	0.06	0.00	0.52	0.00	0.48	0.00	0.35		
12E					20	0.02	0.00	0.51	0.00	0.44	0.00	0.24		
13E				10	5	0.13	0.07	0.22	0.04	0.23	0.02	0.22		
14E					10	0.06	0.01	0.24	0.01	0.25	0.00	0.21		
15E					20	0.02	0.00	0.33	0.00	0.30	0.00	0.18		
16E				20	5	0.08	0.06	0.09	0.05	0.10	0.04	0.10		
17E					10	0.06	0.03	0.10	0.02	0.11	0.01	0.10		
18E					20	0.03	0.01	0.15	0.00	0.14	0.00	0.12		

[z] Appropriate for analyzing data from an RCBD in which multiple trees in a group are assigned the same treatment within each block

[y] Appropriate for analyzing data from a generalized RCBD in which multiple trees are independently assigned the same treatment within each block

to declare significance. Because type I and type II errors are inversely related for given sample size and treatment effect, an increase in the type I error threshold, say to P=0.2, would result in a corresponding increase in power. A strong argument could be made for acceptance of a much higher Type I error rate (probability that treatments responses are declared different when differences result from random factors) in many field trials (4), especially when there are no significant cost differences or potential disadvantages to recommendation based on the "wrong" conclusion. A subset of the analyses presented here were run using a type I error rate of P=0.20: as expected, the power for each test was greatly increased, but the relative value of each experimental design was unaffected (data not shown), with 1 to 2 trees per treatment per block providing the greatest power.

This analysis focused solely on the randomized-complete-block design. When very small field experiments are conducted on uniform sites with no obvious blocking factor, it may be best to avoid use of blocks altogether (5). However, this study suggests that even in the case of significant blocking factor effects, keeping block sizes to the minimum to accommodate all treatments (i.e. using a complete block design) provides the highest power. This becomes even more critical when the effect of the treatment is confounded with block effects. This may happen for example with perennial crops where even small differences in site or plant characteristics can have a cumulative effect over time (5), resulting in responses that display significant treatment-by-block interactions. Randomized-block designs using within-block-replication are especially efficient when each block is fairly homogeneous (2). When there is substantial heterogeneity within blocks, reblocking to smaller blocks seems appropriate or more sophisticated designs such as Latin Square or lattice layouts used. An alternative approach is the use of a linear analysis model with the blocking factor directly incorporated. Both alternatives have the potential to increase hypothesis test power (2) but these approaches may be difficult to implement for many field trials using established perennial crops.

Using this information. In this study we have compared the power of the statistical test for treatment effects under the constraint of fixed total size of the experiment. This has allowed direct comparison of design effects without confounding effects. These analyses pointedly demonstrate that designs in which individual trees are the experimental units are almost always superior to pooling data across several trees when the same number of trees is used per treatment. Our results also suggest that there is no significant advantage to using an experimental design with more than 2 experimental units per treatment per block. Even in the situation where blocking is important and there is a treatment-by-block interaction, only the minimal replication needed to allow estimation of both treatment-by-block and pure error effects is needed (Table 7). However, where there is no treatment-by-block interaction, and a blocking factor with substantial effect can be used in blocking, considerable power is lost by not maximizing the total number of blocks used. As the influence of the blocking factor on the response variable increases, the power advantage of more blocks is substantially increased.

The experimenter's knowledge or expectations of the blocking factor effect and likely blocking factor by treatment interaction should be used to select the best experimental design. If these factors are totally unknown to the experimenter, it may be best to use two experimental units per treatment per block. Clearly the presence of a blocking factor by treatment interaction greatly increases the power of this design *vis-à-vis* a single experimental unit per treatment per replication and only modest power losses would be realized by this choice if there is no interaction.

Certainly there may be advantages to increasing the total number of trees on which data are collected by choosing to sample multiple trees using a similar or even higher number of independent experimental units. In

Table 8. Simulation results with ordered assignment to simulate spatial blocking, in which consecutive trees form multi-tree experimental units and neighboring trees are more like each other than more distant trees in the same block. No experimental increment (treatment has no effect on the response variable) is compared with no or modest block x treatment interaction (α1=0, 0.5). A small experimental increment (treatment incrases response variable by 5%) is compared with no, modest, or large block x treatment interaction (α1=0, 0.5, 1.0). There is no treatment effecton slope between blocking factor and response variable (γ_1=0).

					12 blocks of 1 tree / treatment	6 blocks of 2 trees / treatment		4 blocks of 3 trees / treatment		2 blocks of 6 trees / treatment	
					P of rejecting null hypothesis						
Trial	Slope of response variable vs. covariate (α1)	Treatment effect on response variable (γ0)	Standard error of response variable (se)	Standard error of covariate (sx)	Block × trt as error for F test	Block × trt as error for F test[z]	MSE as error for F test[y]	Block × trt as error for F test[z]	MSE as error for F test[y]	Block × trt as error for F test[z]	MSE as error for F test[y]
1F	0	0%	5	5	0.05	0.05	0.049	0.05	0.05	0.05	0.05
2F				10	0.05	0.05	0.046	0.05	0.06	0.05	0.04
3F				20	0.05	0.05	0.05	0.05	0.05	0.06	0.05
4F			10	5	0.05	0.05	0.05	0.05	0.05	0.05	0.05
5F				10	0.05	0.05	0.05	0.05	0.05	0.05	0.05
6F				20	0.05	0.05	0.05	0.05	0.05	0.05	0.05
10F	0.5	0%	5	5	0.05	0.06	0.08	0.05	0.13	0.02	0.47
11F				10	0.06	0.05	0.14	0.07	0.29	0.12	0.57
12F				20	0.05	0.05	0.26	0.09	0.45	0.12	0.59
13F			10	5	0.05	0.05	0.06	0.05	0.07	0.06	0.21
14F				10	0.05	0.05	0.07	0.05	0.13	0.09	0.48
15F				20	0.05	0.05	0.13	0.07	0.30	0.11	0.57
16F	0	5%	5	5	0.62	0.51	0.63	0.39	0.64	0.15	0.64
17F				10	0.62	0.51	0.61	0.40	0.63	0.15	0.66
18F				20	0.60	0.51	0.62	0.39	0.64	0.16	0.65
19F			10	5	0.20	0.18	0.21	0.14	0.21	0.08	0.21
20F				10	0.20	0.17	0.20	0.14	0.22	0.09	0.21
21F				20	0.20	0.17	0.21	0.14	0.21	0.09	0.22
22F	0.5	5%	5	5	0.58	0.41	0.58	0.22	0.55	0.09	0.54
23F				10	0.52	0.26	0.52	0.12	0.52	0.11	0.67
24F				20	0.37	0.11	0.44	0.08	0.50	0.13	0.64
25F			10	5	0.20	0.18	0.21	0.12	0.22	0.07	0.32
26F				10	0.19	0.14	0.22	0.09	0.27	0.09	0.47
27F				20	0.18	0.10	0.24	0.08	0.36	0.11	0.61
28F	1	5%	5	5	0.52	0.27	0.54	0.11	0.52	0.11	0.67
29F				10	0.36	0.12	0.45	0.08	0.49	0.12	0.64
30F				20	0.18	0.07	0.40	0.10	0.52	0.13	0.62
31F			10	5	0.20	0.14	0.23	0.10	0.27	0.09	0.49
32F				10	0.17	0.10	0.26	0.07	0.36	0.12	0.60
33F				20	0.13	0.06	0.30	0.09	0.47	0.12	0.59

[z] Appropriate for analyzing data from an RCBD in which multiple trees in a group are assigned the same treatment within each block

y Appropriate for analyzing data from a generalized RCBD in which multiple trees are independently assigned the same treatment within each block; in this case it is the wrong error term and treats individual samples in multi-tree experimental unit as a pseudoreplicate, spuriously rejecting the null hypothesis at a high rate.

some cases logistics of treatment application or plot characteristics may make it preferable to use multiple tree experimental units for more accurate application, avoidance of near-est neighbor effects, or efficient use of limited trees with buffer trees between treatments. These results are not intended to suggest that such uses are not appropriate, but rather that

there is no inherent advantage to use of multi-tree blocks to buffer treatment variability.

A survey of the ecological literature revealed that a large proportion of the published studies used pseudoreplication (1). It is possible that some horticultural researchers who routinely use multi-tree experimental units in field trials have found that such designs are more likely to successfully distinguish treatment differences because they are treating each tree as a pseudoreplicate. The analyses presented here underscore that such practices are not only contrary to sound statistical analysis, but may also result in misleading and unsound conclusions.

Literature Cited

1. Heffner, R.A., M.J. Butler and C.K. Reilly. 1996. Pseudoreplication revisited. Ecology 77:2558-2562.
2. Lin, C.S., M.R. Binns, H.D. Voldeng and R. Guillemette. 1993. Performance of randomized block designs in field experiments. Agron. J. 85:168-171.
3. Marini, R.P. 1985. Sample size estimates for peach tree growth and yield experiments. J. Amer. Soc. Hort. Sci. 110:604-608.
4. Marini, R.P. 1999. Are nonsignificant differences really not significant? HortScience 34:761-762.
5. Pearce, S.C. 1995. Some design problems in crop experimentation. I. The use of blocks. Expl. Agric. 31:191-203.
6. Schabenberger, O. 2010. The difference between the GRBD and the RCBD with subsampling. http://www.stat.org.vt.edu/schabenb/teach/GRB-DvsRCBD.pdf <<Accessed 17 May 2010>>

CALL FOR WILDER SILVER MEDAL NOMINATIONS

The Wilder Committee of the American Pomological Society (APS) invites nominations for the 2011 Wilder Silver Medal Award. All active members of APS are eligible to submit nominations. The award was established in 1873 in honor of Marshall P. Wilder, the founder and first president of APS. The award consists of a beautifully engraved medal which is presented to the recipient at the annual meeting of APS, held during the American Society for Horticultural Science annual meeting.

The Wilder Medal is presented to individuals or organizations that have rendered outstanding service to horticulture in the area of pomology. Special consideration is given to work relating to the origination and introduction of meritorious fruit cultivars. Individuals associated with either commercial concerns or professional organizations will be considered if their introductions are truly superior and have been widely planted. Significant contributions to the science and practice of pomology other than through fruit breeding will also be considered. Such contributions may relate to any important area of fruit production such as rootstock development and evaluation, anatomical and morphological studies, or noteworthy publications in any of the above subjects. Information about the award, past recipients, etc. can be found on the APS web site at http://americanpomological.org/wilder1.html.

To obtain nomination guidelines, please contact committee chairperson:
Dr. Douglas Archbold, Department of Horticulture, University of Kentucky
Phone: 859-257-3352; fax: 859-257-2589; e-mail: darchbol@uky.edu

Nominations must be submitted by May 2, 2011.