Illustrating Split-Plot Designs With 3-D Models

James A. Alloway, Jr.
EMSQ Associates, 152 Grandview Avenue, Catskill, NY 12414
jalloway@TouchAndSeeDOE.com

Abstract
While split-plot designs have been available for nearly 90 years, interest in this type of design has spiked during the last 20 years. Most statistical software packages used in quality improvement efforts have either recently added or now emphasize this design. Although software handles the increased computational demands of the analysis, the experimental team must recognize when a split-plot design is beneficial and also be aware that the analysis must reflect any restrictions on randomization during data collection. Split-plot designs address a key objection to utilizing statistically designed experiments: complete randomization, which increases the time and effort to conduct experiments. If practitioners aren't comfortable with these designs, they will not use them. We need to simplify the presentation of this family of experiments for quality practitioners who lack extensive backgrounds in statistical methods. One approach is to utilize 3-D models to explore the structural differences between completely randomized factorial designs and factorial designs with restrictions on randomization. Once this concept is clear, it is easier to plan for data collection, understand the software output, and explain the results to others.

Key words: Split-plot design, DOE, 3-D models, Pedagogy

1.Split-Plot Experiments

Split-plot experiments have received increasing coverage in the literature during the past 20 years. See for example, Box (1996), Dowalski, et al. (2007), Jones and Nachtsheim (2009), Kowalski and Potcner (2003), Potcner and Kowalski (2004), Robinson, et al. (2009), and Vining (2009). They are a type of factorial design used when some factor levels are inconvenient, time consuming, or more difficult to change than others. Total experimental time and costs are reduced by not having to change factor levels as often as with a completely randomized design.

Analytical software packages used in Six Sigma deployments, such as *Design-Expert*, *JMP*, and *Minitab*, have added split-plot designs in recent releases, making this design family readily available to practitioners. Other than identifying which factors are hard to change, no knowledge of the structure or analysis of split-plot designs is required to use the software.

Increased awareness through publications and simplification of calculations via software are necessary, but not complete conditions for the successful application of split-plot designs. All experimenters must understand the fundamentals of how these designs are created and analyzed so that they can successfully lead problem solving teams and explain results to interested parties. Without this knowledge, practitioners may misapply the technique and misinterpret the software results.

This work utilizes 3-dimensional models of the factor space and the response space to make split-plot design concepts and basic calculations easier to understand by non statisticians. We begin with a notation for split-plot designs and then present five designs, each utilizing 16 runs, to illustrate the basic properties of split-plot designs.

2. Split-Plot Design Notation

The standard $2_R^{k\text{-}p}$ factorial notation is well known within the DOE community, see Box, Hunter, and Hunter (2005), and provides a wealth of information about the design:

Number of factor levels:	2
Number of factors:	k
Design resolution:	R
Number of runs	$2^{(k-p)}$
Fraction of full design:	2^{-p}

The utilization of a similar notation for split-plot designs is lacking in much of the DOE literature, with Bisgaard (2000) among the exceptions. A standard notation to fully describe a split-plot design is beneficial when comparing and contrasting the different design configurations.

The proposed notation is shown in Equation 1. As with standard factorial notation, the *2* refers to the number of factor levels and the subscript *R* refers to the design resolution. In the exponent, *k* refers to the number of factors and *p* is used for fractions in the whole plot and subplot. The subscripts *w* and *s* refer respectively to the whole plot and subplot components.

$$2_R^{(k_w-p_w)+(k_s-p_s)} \qquad \textbf{(1)}$$

By inserting the values for each *k* and *p* into the equation, one can readily determine the following:

Total number of runs = $2^{(k_w-p_w)+(k_s-p_s)}$

Number of subplots = 2^{k_w}

Number of separate randomizations = $2^{k_w}+1$

This notation makes the number of factors that are hard to change explicit. It accommodates full and fractional designs at both the whole plot and split-plot levels. The notation highlights that the allocation of factors to whole plots and split-plots is not commutative. That is, a $2^{(1)+(2)}$ design is not the same as a $2^{(2)+(1)}$ design.

The literature frequently states that split-plot designs require two randomizations and the reader is left to infer that one randomization is for the whole plot and the other is for the subplot. This information is sufficient to understand the reason for two error terms in the ANOVA calculations, but does not address what happens when one actually plans the experiment. *Each* subplot requires a unique randomization, so this disconnect may be a source of confusion for new users of split-plot designs.

3. Visualization to Help Improve Understanding

Most of the split-plot literature assumes that the reader has at least an intermediate statistical education. Many Six Sigma practitioners are not at this level, so they benefit from a non mathematical approach to understand the core concepts.

The remaining sections illustrate the fundamental concepts of split-plot designs using physical 3-D models. Physically building the model involves additional senses to aid the learning process, helps the student to visualize how the design is constructed and executed, and lays the foundation for why there are two measures of experimental error in the ANOVA calculations.

To help differentiate split-plot designs from full factorial and blocked factorial designs, each of the following cases is accompanied by a figure depicting the random run order for that design. *Minitab* version 17 was used to create the random run order, but any other software package will give equivalent results.

3.1 Color Code Key for 3-D Models

In the following figures, the corner nodes and connecting struts are color coded based on their role. Not all colors are used in each figure.

Element/color	Role
Black node	Whole plot experimental unit
White node	Whole plot observational unit
Black strut	Whole plot structure
White strut	Links whole plot experimental unit to whole plot observational unit
Yellow strut	Links whole plot observational unit to subplot observational units
Grey node	Subplot experimental unit and observational unit
Blue strut	Subplot structure
Blue node	Structural connector: no role in the statistical design

Table 1. Legend for node and strut colors in the figures.

3.2 The 2^4 Design

Figure 1 shows the 3-D model for this base design: two cubes with eight nodes for a total of $2^4 = 16$ nodes, where each node represents an experimental run. Each node is both an experimental unit and an observational unit.

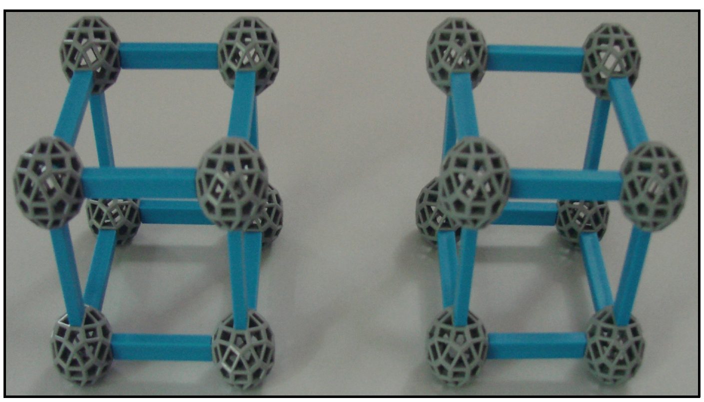

Figure 1. 3-D model for full factorial design.

Figure 2 shows the random run order for this experiment. All 16 values are allocated between the two cubes, indicating a single randomization. This single randomization results in a single estimate for experimental error.

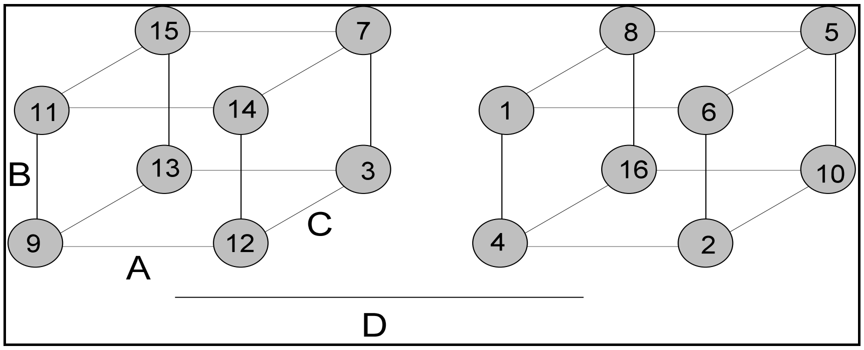

Figure 2. Random run order for the full factorial design.

3.3 The 2^4 Design in Blocks

Split-plot designs are sometimes described as a factorial design run in blocks. However, a blocked design is not necessarily a split-plot design. It is important for those new to split-plot designs to recognize this difference.

To illustrate, Figure 3 shows a four-factor design run in two blocks. There are four cubes, with the top two representing block one and the bottom two representing block two. One block is randomly selected and all runs within that block are completed before switching to the second block. In this model, each corner point does not necessarily represent an experimental run. Recall from Table 1 that the grey nodes are the both experimental units and observational units. The blue nodes are structural connectors for the physical model.

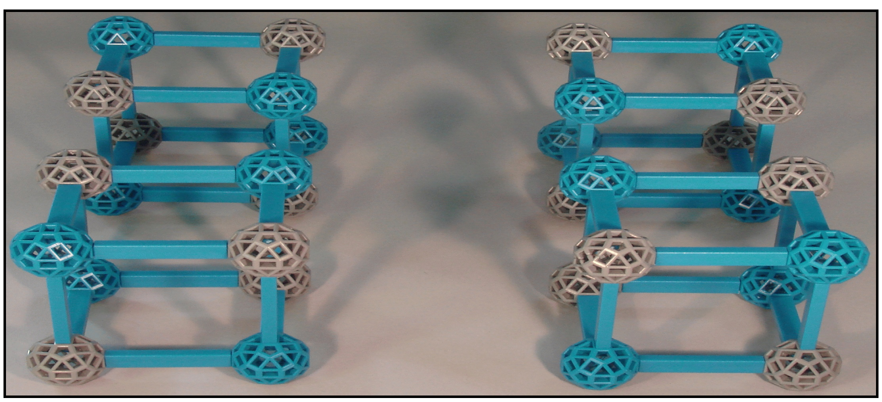

Figure 3. Full factorial design run in two blocks. Block one consists of the top two cubes, block two consists of the bottom two cubes.

Using standard factorial notation, there are $2^4 = 16$ total runs, with eight runs per block. Unlike the single randomization of the full 2^4 design, each block requires a unique randomization for run order, as shown in Figure 4. Note that the values 1- 8 appear in each block. All factors are varied without respect for difficulty in changing levels in the blocked design.

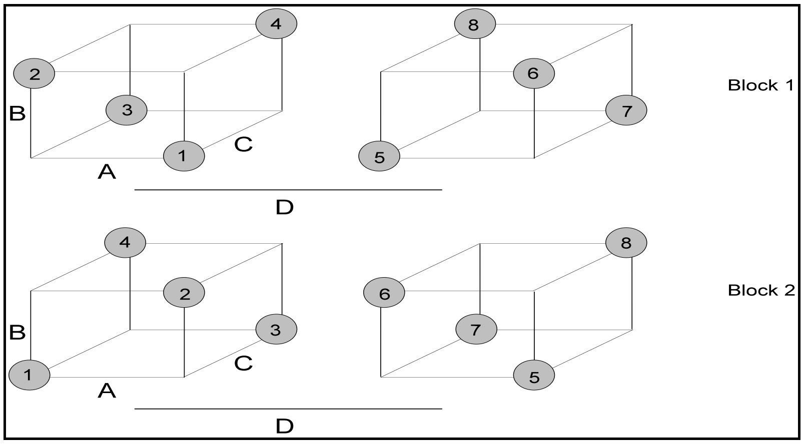

Figure 4. Random run order for blocked design.

3.4 Summary of the Basic Cases

In the full factorial design, all experimental treatments are run in each replicate. In the present case, there was a single replicate. All factor effects can be determined free of confounding. There are no restrictions on the random order of the runs; thus a single randomization sequence suffices, resulting in a single estimate for experimental error.

In a blocked design, the blocking variable is not a factor in the experiment and is beyond the experimenter's control. To maintain 16 runs for consistent comparison with the other cases, the blocking factor could be considered a potential fifth factor. Blocking reduces experimental error by making the blocks as homogeneous as possible. Every treatment combination does not appear in each block, but there is a symmetry and balance between the blocks. In the case of two blocks, the highest order interaction, in this case the ABCD interaction, is confounded with blocks. A design with blocks requires each block to be randomized independently with all factors (but not all factor levels) appearing in each block.

3.5 The $2^{(1)+(3)}$ Design

We now examine three variants of a 16 run split-plot design, beginning with a single hard to change factor and three easy to change factors. Using the proposed split-plot notation, there are $2^{(1)+(3)} = 2^4 = 16$ runs, $2^{(1)} = 2$ subplots, and $2^{(1)} + 1 = 3$ separate randomizations: one for the whole plot and one for each subplot.

Physically building the model in Figure 5 helps to visualize the split-plot design structure and run order. The model contains attributes of both the full factorial and the blocked designs from the previous two sections.

We begin with the whole plot (the black nodes and struts at the bottom of the figure), since it literally supports the subplots, represented by the two cubes at the top of Figure 5. The grey nodes represent both the experimental unit and the observational unit. The white nodes represent the whole plot observational unit and connect to the black nodes (whole plot experimental units) with a white strut. The yellow struts show the connection between the whole plot observational unit and the subplot observational units.

The whole plot portion of the design is similar, but not identical to the blocks in the priordesign. Randomization is restricted within the subplots of each whole plot treatment.

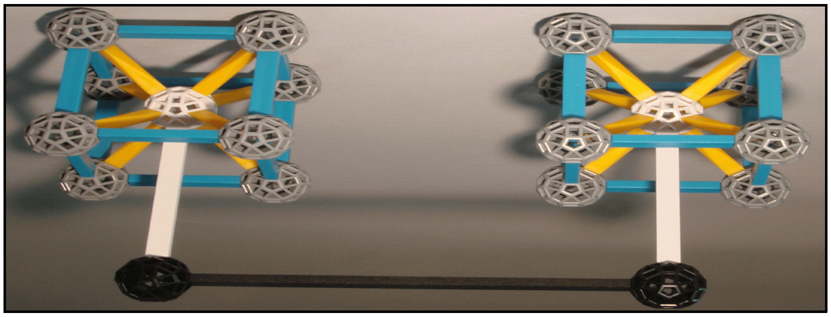

Figure 5. Split-plot design with one hard to change factor and three easy to change factors.

That is, one whole plot treatment is selected and all subplot treatments within that treatment are run before moving to the next whole plot treatment. Unlike the blocked design, the whole plot factor is under experimenter control. Also, the highest order interaction term is not confounded. Each subplot factor within each whole plot combination appears at all levels, as contrasted with the blocked design, where not every combination appeared in each block. The whole plot nodes represent experimental units, the physical unit to which the experimental treatment is applied.

The subplot behaves like the full factorial design. All subplot factors appear at all levels and there is no restriction on randomization within each subplot.

Figure 6 shows the random order for this experiment. One of the whole plot levels is randomly selected and then all points in the corresponding subplot are run. Each subplot has a separate random run order, resulting in a total of three randomizations for this experiment.

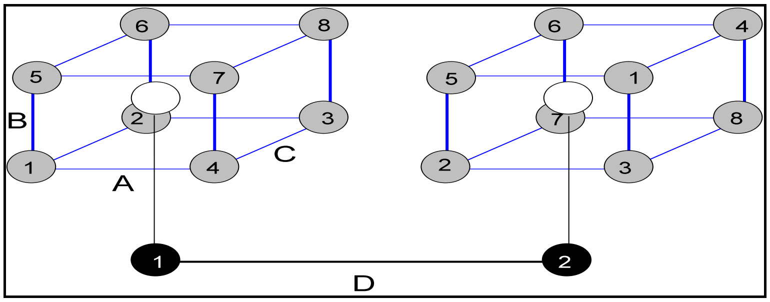

Figure 6. Random run order for $2^{(1)+(3)}$ design.

3.6 The $2^{(2)+(2)}$ Design

This design contains two hard to control factors in the whole plot and two easy to control factors in the subplots. In this design, there are $2^{(2)+(2)} = 2^4 = 16$ runs, $2^{(2)} = 4$ subplots, and $2^{(2)} + 1 = 5$ separate randomizations; one for the whole plot, and one for each of the four subplots.

We begin by building the two factor whole plot square (in black) at the bottom of Figure 7, then add the two factor subplots (also squares) at each whole plot treatment. The run order of the four whole plot combinations is determined first, then a separate run order for each corresponding subplot is determined, as shown in Figure 8.

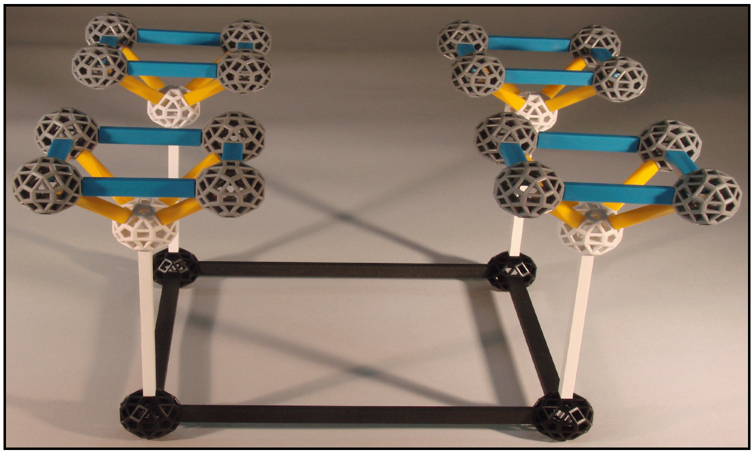

Figure 7. 3-D model for the $2^{(2)+(2)}$ design.

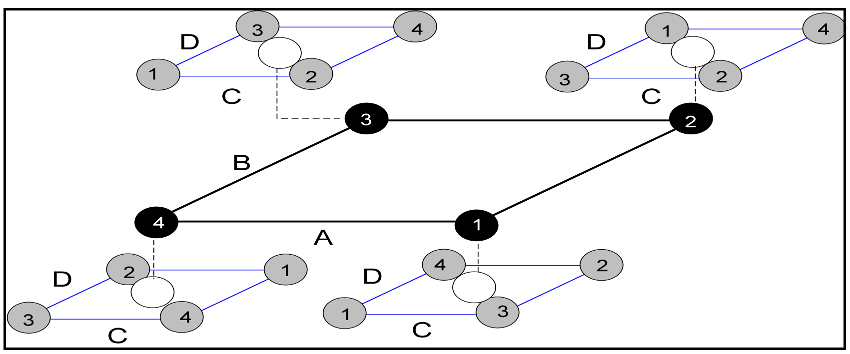

Figure 8. Random run sequence for the $2^{(2)+(2)}$ design.

3.7 The $2^{(3)+(1)}$ Design

The final case considers three hard to change factors in the whole plot and one easy to change factor in the subplot. There are $2^{(3)+(1)} = 2^4 = 16$ runs, $2^{(3)} = 8$ subplots, and $2^{(3)}+1 = 9$ separate randomizations for run order; one for the whole plot, and one for each of the eight subplots.

As with the other split-plot designs, we begin with the whole plot factors to create the black cube in Figure 9. Each cube corner point supports a single factor subplot, which are shown as blue struts with grey nodes.

Figure 10 shows the run order for this experiment. Each whole plot is selected in the order indicated, then the corresponding subplot is run before moving onto the next whole plot.

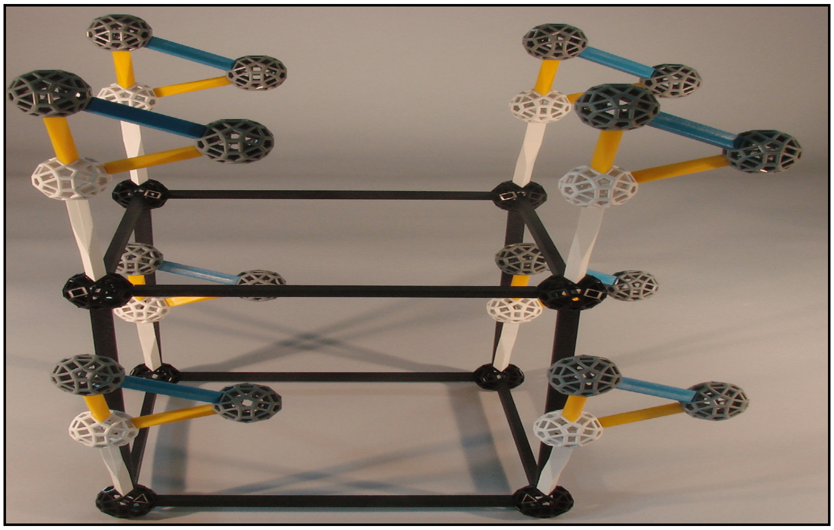

Figure 9. 3-D model of the $2^{(3)+(1)}$ design.

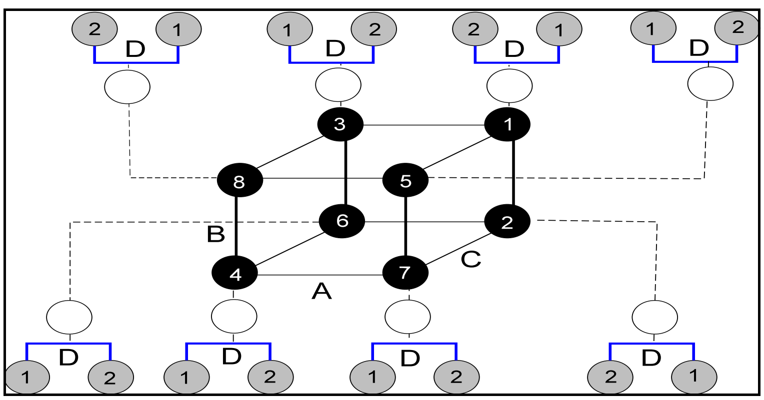

Figure 10. Random run order for the $2^{(3)+(1)}$ design.

3.8 Split-Plot 3-D Model Summary

The 3-D models in the previous three sections help illustrate why a standard split-plot notation is essential for effective communication. All three cases utilized 16 runs, but the allocation of factors between the whole plots and split plots resulted in vastly different structures. The models clearly demonstrate that a $2^{(1)+(2)}$ design is not the same as a $2^{(2)+(1)}$ design.

In split-plot designs, the method used to build the 3-D models mimics how the actual experiment is run. The whole plot experimental units come first. In the physical model, they serve as the support base for the subplot components. In an experiment, the whole plot treatments are applied to the experimental units first. Next the subplot treatments are applied to each whole plot experimental unit. Randomization takes place first at the whole plot level, and then individually within each subplot.

4. Response Space Calculations

Thus far, the 3-D models represented the design factor space. The response space of the design is represented with white struts with a scale on them. Figure 11 shows the scaled white struts added to the $2^{(2)+(2)}$ design. For clarity, these new struts are shown for only one subplot.

For subplots, the response is measured directly at the grey nodes. Recall that the observational and experimental units are the same in the subplot, just as in the standard factorial design. The subplot effects are calculated as the difference between the average high and low level responses for the factors.

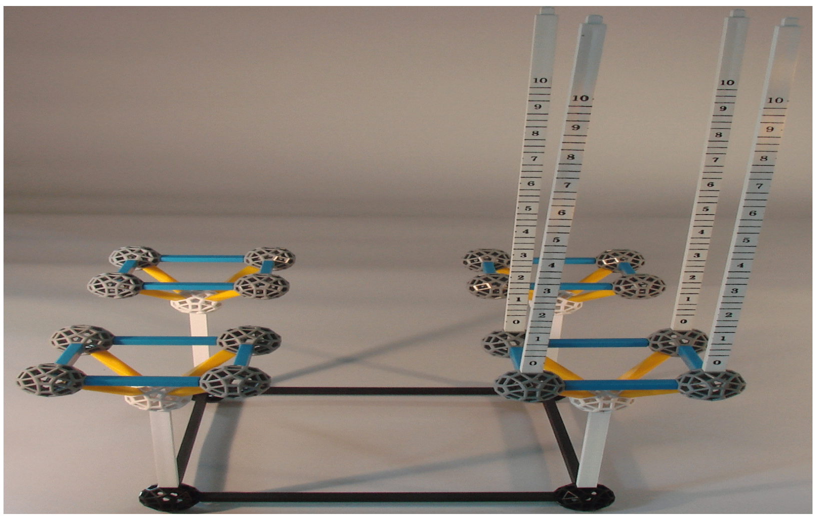

Figure 11. Response space for a single subplot in a $2^{(1)+(2)}$ design.

The whole plot effect calculations are more involved. The whole plot response space is shown in Figure 12, with white scaled struts attached to the white nodes (observational units) of the whole plot. A plain white strut connects the black node to the white node, linking the experimental unit to the observational unit. Unlike the factorial design, the whole plot observational units (white nodes) are not the same as the experimental units (black nodes).

The whole plot observational unit is connected by yellow struts to the subplot observational/experimental units. The response at each whole plot treatment is the average of the connected subplot responses. These values are then used in the traditional way to determine the whole plot effects.

The effect calculations for the subplot and whole plot effects rely on different randomization schemes, different methods of calculation, and thus different degrees of freedom. The result is two different estimates of experimental error, one for the whole plot and another for the subplot.

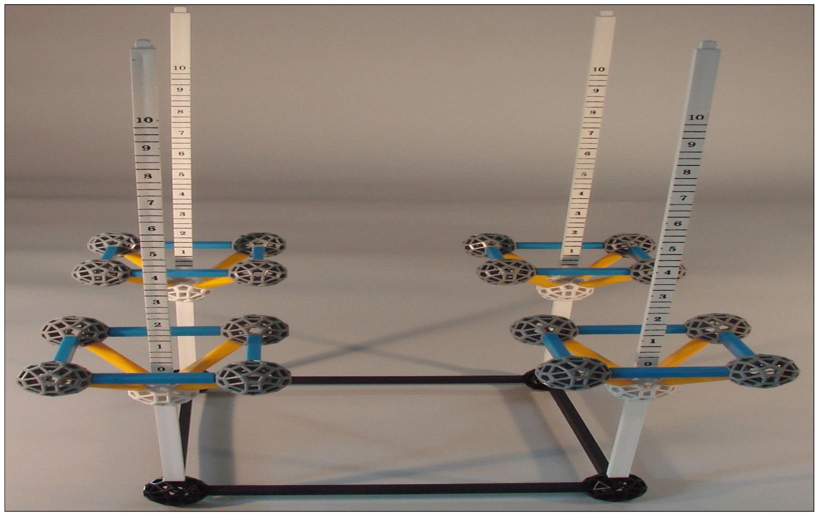

Figure 12. Response space for the whole plot of a $2^{(1)+(2)}$ design.

5. Summary

3-D models provide students with insights into the structure and analysis of split-plot designs. They provided a readily understandable framework to compare and contrast split-plot designs with the more familiar full factorial and blocked factorial designs. They help highlight the experiment-within-an-experiment structure of split-plot designs by showing that the whole plot portion separates the experimental unit from the observational unit. They are especially helpful in distinguishing between experimental and observational units, and how that difference affects the ANOVA calculations in split-plot designs.

References

Bisgaard, Soren. (2000). The Design and Analysis of 2k-p x 2q-r Split Plot Experiments. *Journal of Quality Technology*, 32 (1): 39 - 56.

Box, George. (1996). Quality Quandaries, Split Plot Experiments. *Quality Engineering*. 8 (3): 515-520. Marcel Dekker, Inc.

Box, George E. P., J. Stuart Hunter, and William G. Hunter. (2005). *Statistics for Experimenters*. Hoboken, NJ. Wiler-Interscience.

Dowalski, Scott M., Peter A. Parker, G., and Goeffrey Vining. (2007). Tutorial: Industrial Split-plot Experiments. *Quality Engineering*. 19:1-15.

Jones, Bradley, Christopher J. Nachtsheim. (2009). Split-Plot Designs: What, Why and How. *Journal of Quality Technology* 41(4): 340 - 361.

Kowalski, Scott M., Kevin J. Potcner. (2003). How to Recognize a Split-Plot Experiment. *Quality Progress*. November. ASQ.

Potcner, Kevin J., Scott M. Kowalski. (2004). How to Analyze a Split-Plot Experiment. *Quality Progress*. December. ASQ.

Robinson, Timothy J., William B. Brenneman, William R. Myers. (2009). An Intuitive Graphical Approach to Understanding the Split-Plot Experiment. *Journal of Statistics Education*. 17:1.

Vining, Geoff. (2009). Technical Advice: Industrial Split-Plot Experiments. *Quality Engineering.* 22(1): 54-57.