Designing Experiments that Combine Mixture Components with Process Factors

Apply powerful statistical tools to optimize your formula while simultaneously finding the peak process parameters.

Mark J. Anderson* and Patrick J. Whitcomb

Stat-Ease, Inc.

2021 East Hennepin Ave, Suite 480, Minneapolis, MN 55413

Telephone: 612.378.9449, Fax: 612.378.2152

***E-mail: Mark@StatEase.com**

The typical strategy for design of experiments (DOE) in the chemical process industry is:

1. Fine tune the formulation via mixture design[1]
2. Optimize the process with factorial design and response surface methods[2]

To keep things simple, these two steps are usually handled separately by the chemist and chemical engineer; respectively. However, interactions between compositional variables and process factors cannot be revealed by this simplistic approach. In this article we show you how to do a comprehensive experiment that combines mixture components with process factors in one "crossed" design.

A relatively simple case study

To illustrate how to do a crossed mixture-process design, we present a relatively simple case study from "Experiments with Mixtures" by Cornell.[3] (This textbook provides a wealth of statistical detail on design of experiments for mixtures, including the crossing of process factors.) The case study involves three vinyl plasticizers (X_1, X_2, X_3) processed at two levels each for extrusion rate (Z_1) and drying temperature (Z_2). Many other components go into vinyl formulation (stabilizers, lubricants, drying agents and resins), but the percentages of all non-plasticizer ingredients were held fixed throughout the experiment, so they are non-factors.

Figure 1 shows a picture of the crossed design. The triangles represent the mixtures, which must be repeated at the four combinations of the process factors (Z_1 and Z_2).

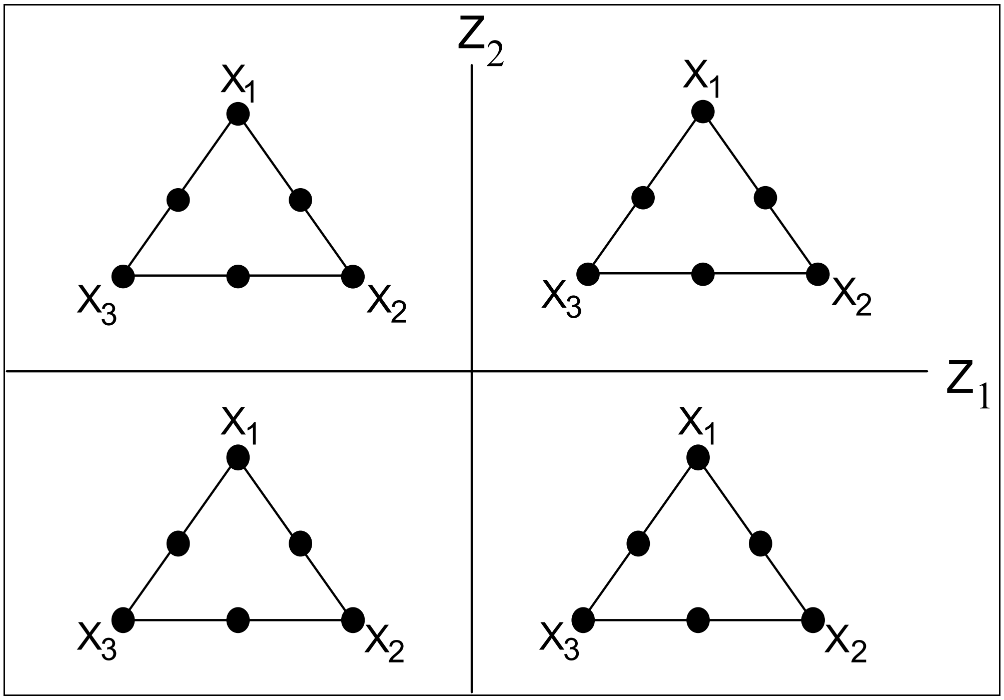

Figure 1. Crossed Mixture-Process Design

The scale on the mixture diagrams go from zero to one based on relative proportions of the three ingredients. The vertices represent pure component blends (X_1, X_2 and X_3). Binary blends (0.5/0.5 combinations of any two plasticizers) occur at the midpoints of the sides on the triangle. The interior space, empty in this case, represent three part blends. This points on the mixture plots come from an standard experimental design called a "simplex lattice." These designs can be tailored for the degree of polynomial you want to fit: linear (1st degree), quadratic (2nd degree) or cubic (3rd degree). The experiment on the vinyl formulation was done with a second degree simplex lattice, thus revealing any two-component interactions between plasticizers. If you choose this design for your experimental work, we recommend you augment it with a three-part blend called a "centroid." Results from this blend will reveal potential problems with lack of fit in the quadratic mode.

Table 1 shows the experimental design in terms of coded factor levels:

- For mixture components from 0 to 1 for least to most, respectively
- For process factors from –1 to +1 for lowest to highest levels, respectively

For proprietary reasons, the experimenters did not reveal the actual units of measure for the variables, but this does not matter for illustration purposes. All calculations will done in coded form in any case. The experimenters measured the effects of these variables on thickness of an automotive seat cover. Again for proprietary reasons, the results have been re-scaled, so no units of measure are reported, but the relative results remain relevant. The entire run was replicated to gain statistical power, so there's a total of 48 runs (6 blends at 4 process combinations done 2 times). To simplify tabulation the replicates are shown as two columns, and each mixture-process combination shown as a unique row listed in a standard order. However, the actual experiment was completely randomized to insure against lurking factors such as material degradation, machine wear, ambient changes and the like. Randomization is an essential element of good statistical design.

Creating a mathematical model

The experimenters wanted better control of the thickness response. The desired results depended on the model of automobile. For example, thicker vinyl seats might be needed for a pickup truck aimed at the rugged outdoors type. On the other hand, a thinner cover might be needed to cut costs on the low-end "econocar." The outcome of a statistically significant DOE is a polynomial model that can be used to predict the response at any combination of tested variables. As you can see from the derivation below, the models for crossed mixture-process designs can be very cumbersome, even for a relatively simple study like the one done on the vinyl seat-covers. In this case, crossing the six term mixture model

(1) $Y(x)=\beta_1X_1+\beta_2X_2+\beta_3X_3+\beta_{12}X_1X_2+\beta_{13}X_1X_3+\beta_{23}X_2X_3$

with the four term factorial model

(2) $Y(Z)=\alpha_0+\alpha_1Z_1+\alpha_2Z_2+\alpha_{12}Z_1Z_2$

produces a twenty-four term model.

(3) $Y(X,Z)=Y(X)\times Y(Z)=6x4=24\text{ terms}$

The letter Y symbolizes the response. The first equation (with the X-variables) represents the mixture part of the design. The Greek letter beta represents the unknown coefficients. Another article in this series[1] describes how these polynomials are constructed to account for the overall constraint that all mixture components must sum to one. Mixture models can be recognized by their lack of an intercept. The one shown above is second order. The second order terms, such as AB, reveal interactions.

The second equation (comprised of Z-variables) represents the process side of the design. It's sometimes called a "factorial" model. The Greek letter alpha represents the unknown coefficients. For more details on this model and more complex polynomials used for response surface methods, see the first part of this series of articles on DOE.[2]

The fitted equation for the vinyl study is:

Thickness (Y) =
$8.88X_1+6.00X_2+6.50X_3+11.25X_1X_2+5.75X_1X_3+2.00X_2X_3$
$-0.63X_1Z_1+0.00X_2Z_1+1.00X_3Z_1-0.75X_1X_2Z_1-4.25X_1X_3Z_1+1.00X_2X_3Z_1$
$-0.38X_1Z_2+0.75X_2Z_2-0.75X_3Z_2-3.75X_1X_2Z_2-2.25\,X_1X_3Z_2+5.00X_2X_3Z_2$
$-2.38X_1Z_1Z_2-1.25X_2Z_1Z_2-0.25X_3Z_1Z_2-8.75X_1X_2Z_1Z_2-3.25\,X_1X_3Z_1Z_2-2.00X_2X_3Z_1Z_2$

This predictive model will be used to generate response surface graphs, which make interpretation much easier than looking at all the coefficients. However, for those of you who want to dissect the equation, notice that the first line contains only mixture components (X-variables). It represents the blending properties, averaged over the various process conditions. The second line of the equation reveals the linear effect of first process factor (Z_1), which shifts the mean response at any given combination of mixture components. The third line shows the linear effect of the second process factor (Z_2). The last line of the equation represents interactions between process factors and the mixture. When these complex interactions are present, you will see the shape of the response surface to change as process conditions are varied.

Analysis of variance (ANOVA) shows the overall equation to be highly significant (Prob>F of <0.0001). However, you will observe that some of the coefficients in the model are at or near zero. These terms could be eliminated via manual reduction or by use of a standard computerized algorithm such as backward stepwise regression. In this case there's no advantage to model reduction because of the presence of statistically-significant higher-order interactions such as $X_1Z_1Z_2$ (Prob>F of <0.01) and $X_1X_2Z_1Z_2$ (Prob>F of <0.01), which must be supported by their lower order 'parents.' (These results justify the application of the crossed design tool because they involve combinations of mixture and process. Interactions such as these would never be revealed by traditional one-factor-at-a-time (OFAT) experiments, or even by more sophisticated DOE's done separately on mixture versus process.) In any case, removal of insignificant terms causes little impact on the predictions or the resulting response surface maps that you will focus on for your reports, so we advise that you work with the complete model.

Using graphs to tell the story

A simple way to interpret the results is to produce contour plots of the thickness versus the composition of the three plasticizers (X_1, X_2, and X_3) at all four combinations of the two process factors – extrusion rate (Z_1) and drying temperature (Z_2). This can be seen in Figure 2.

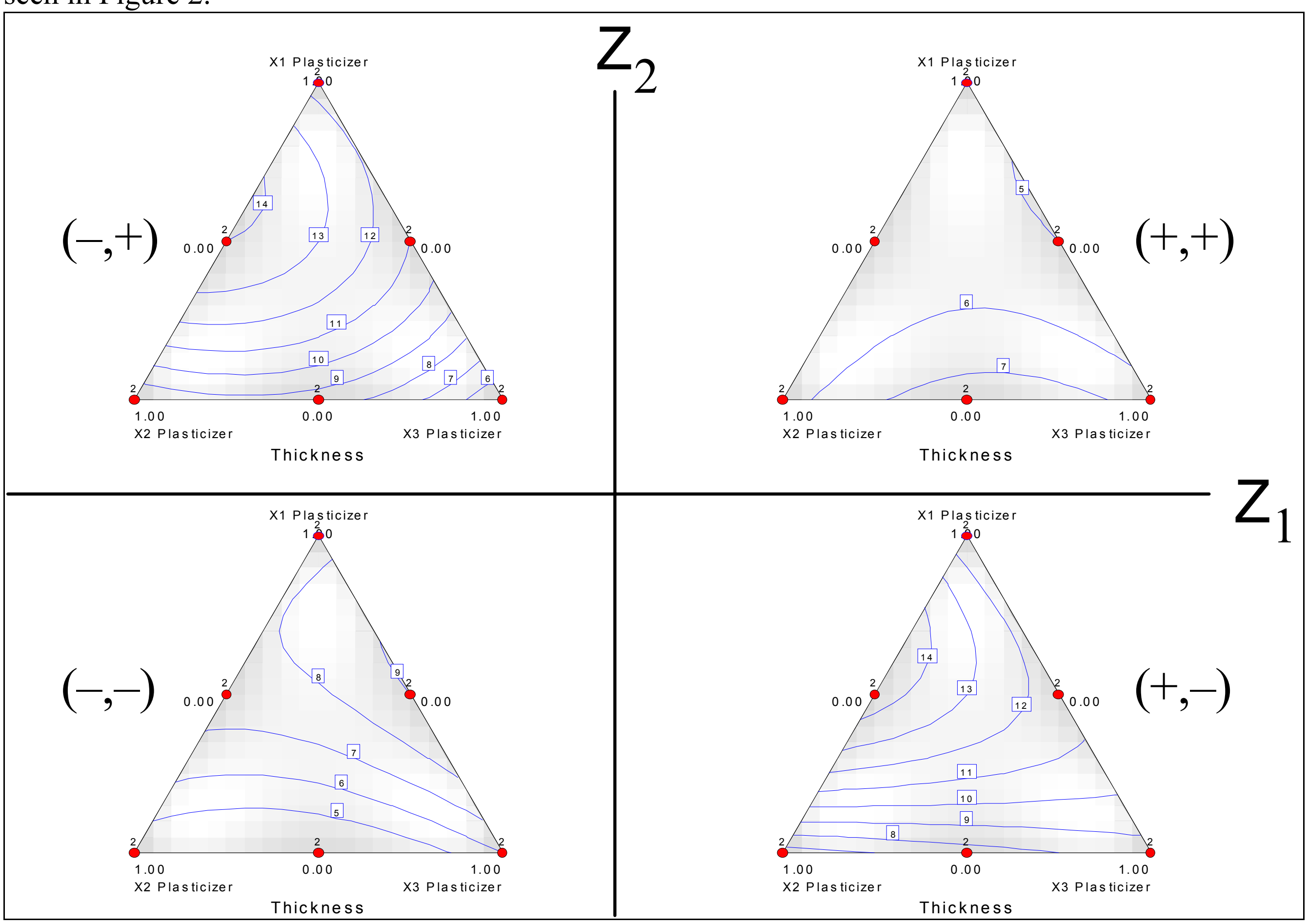

Figure 2. Contour Plots for Mixture-Process Design on Vinyl

Notice how the shape of the contours change as the process conditions vary. From this you now know that the impact of the different plasticizers depends on how the vinyl is processed. But which direction to go will depend on what thickness you desire. Let's assume you want to maximize thickness. In this case you will want to adjust the process conditions to either the high level of extrusion rate with drying temperature low (+, –), or the low rate at the high temperature (–, +). It makes more sense to go for the higher rate for production purposes, which means that you want the lower drying temperature. A 3D version of the contour plot at these conditions is shown in Figure 3.

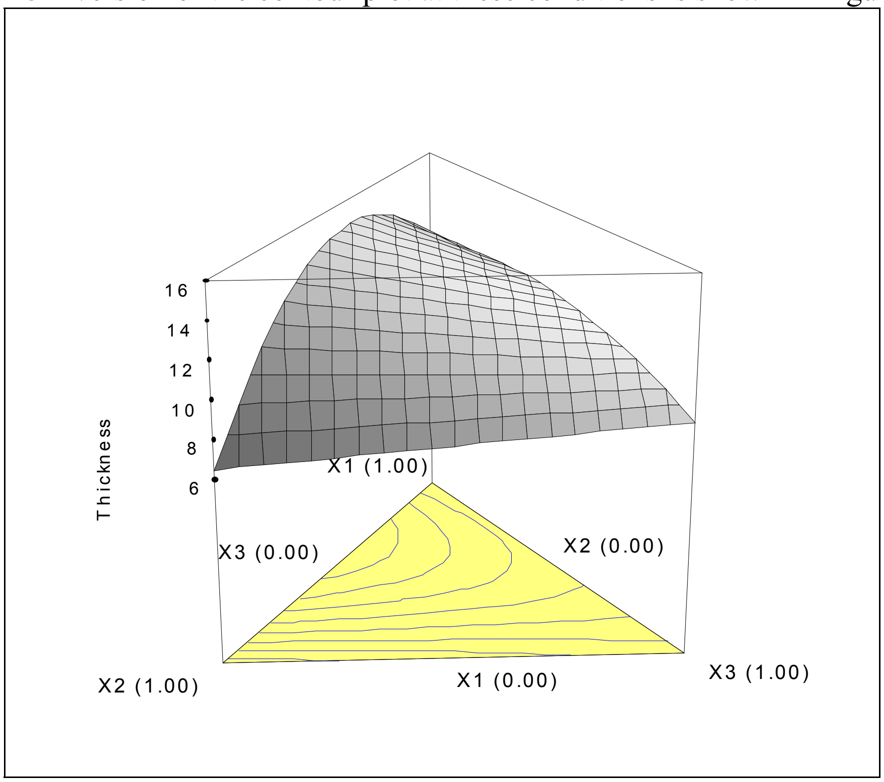

Figure 3. 3D Graph of Mixture at Process Conditions that Produce Thick Vinyl (high extrusion rate, low drying temperature)

The peak thickness occurs when the X_3 plasticizer is excluded from the formulation. The maximum response comes from a binary blend of X_1 and X_2, with somewhat better results with more of the X_1 plasticizer. Numerical optimization, performed on the same DOE software used to fit the data and generate the plots,[4] reveals a peak at the 60/40 blend of X_1/X_2, which produces a (scaled) thickness of 14.7.

Playing the "What-If" game

The solution noted above might work only for certain models of automobiles, such as the heavy-duty trucks. What if a thinner vinyl is needed for the cheaper vehicles? Let's go back to Figure 2 and re-consider our options. Notice that the lower-valued contours occur when you set process conditions to either the low level of extrusion rate with drying temperature low (–, –), or the high rate at the high temperature (+, +). It makes more sense to go for the higher rate for production purposes, which means that you must go to the higher drying temperature. A 3D version of the contour plot at these conditions is shown in Figure 4.

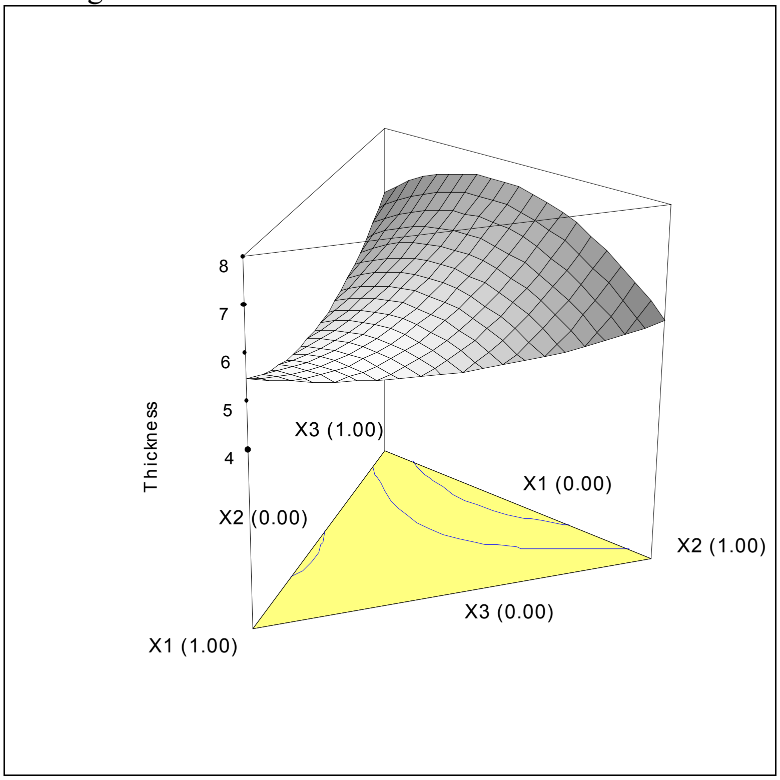

Figure 4. 3D Graph of Mixture at Process Conditions that Produce Thin Vinyl (high extrusion rate, high drying temperature)

At the processing conditions noted above, the vinyl is thinnest at two points:

- A binary blend of X_1 and X_2 plasticizers (along the X_1 - X_2 edge where X_3 is 0)
- A binary blend of X_1 and X_3 plasticizers (along the $X_1 – X_3$ edge where X_2 is 0)

The first option is intriguing because it means that you could use nearly the same blend as that needed for the thick vinyl, but process it differently to make it thin. For example, a 50/50 blend of X_1/X_2 at the high levels of both process factors produces a scaled thickness of 5. By simply lowering drying temperature, the same blend produces a thickness of 14.5. Figure 5 shows the interaction of the process factors at the 50/50 blend of X_1/X_2.

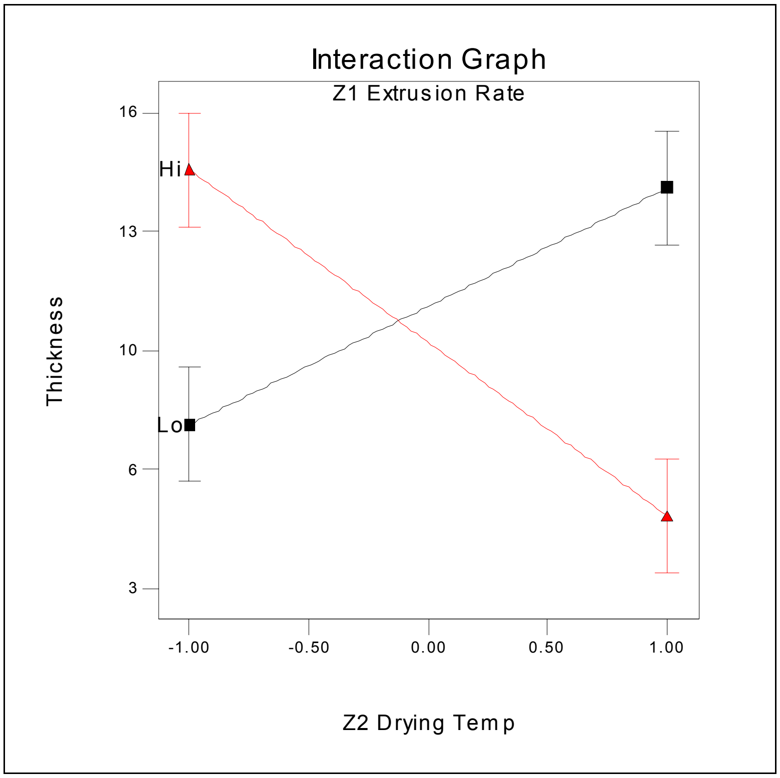

Figure 5: Interaction of Drying Temperature and Extrusion Rate (for 50/50 blend of X_1/X_2 plasticizers)

The results for high extrusion rate (going from upper left to lower right) forms the operating line. To get any thickness from 14.5 down to 5, simply adjust drying temperature from high to low, respectively. However, when making predictions such as this, always remember that actual results may vary due to variations in the blending, the process, the sampling and the testing. In addition, the model itself may be off somewhat because it's based on sample data. For example, for the peal outcome of 14.5, the statistical prediction interval at 95% confidence is 10.67 to 18.33. In other words, when doing confirmation runs, don't be surprised to see individual outcomes somewhat above or below the predictions.

Looking at the wide intervals of prediction may lead to further experimentation aimed at "robust design," which will stabilize the process performance. The DOE arsenal includes a tool called "propagation of error" (or POE) that can be applied for reduction of error transmitted from poorly controlled factors. This and other aspects of robust design, a relatively new arena for DOE, may become the subject of a future article for this series.

More complicated crossed designs

What we've illustrated is a very simple mixture process design that involves only three mixture components and two process factors. Additional mixture and/or process variables will probably push the number of combinations beyond the practical limits of material and time. For example, going to a third process factor on a three-component mixture pushes the total runs to 48. Other complications may arise, such as:

- Constraints on individual components in the mixture.

- Added categorical factors such as who supplies what type of material.

Good DOE software can set up "optimal" designs that minimize the number of experiments regardless of constraints. The last reference provides details on how this can be accomplished.[5]

Conclusion

The case study on the automotive vinyl shows how you can apply advanced tools of DOE to simultaneously optimize your mixture formulation and processing conditions, taking advantage of complex interactions in the system. Response surface graphics, which can be produced with statistical software, make it easy to find the peak performance. If you must juggle many responses to keep your product in specification, numerical optimization approaches are available to manipulate the predictive models and find the "sweet spot" for both mixture and process variables.

Literature Cited

(1) Anderson, M.J., Whitcomb, P.J., "Find the Most Favorable Formulations," *Chemical Engineering Progress,* April 1998.
(2) Anderson, M.J., Whitcomb, P.J., "Optimize Your Process-Optimization Efforts," *Chemical Engineering Progress,* December 1996.
(3) Cornell, *Experiments with Mixtures,* 2nd ed., Example 7-4, John Wiley & Sons, Inc, New York, 1990.
(4) Helseth, et al, *Design-Expert*, Version 6 for Windows, Stat-Ease, Inc, Minneapolis, 1999 ($995).
(5) Anderson, M.J., Whitcomb, P.J., "Computer-Aided Tools for Optimal Mixture Design," *Paint and Coatings Industry,* November 1999.

Table 1. Design Matrix and Data for Vinyl Study

Id[a]	X_1 Plasticizer	X_2 Plasticizer	X_3 Plasticizer	Z_1 Extrusion rate	Z_2 Drying temp	Y Thickness [b] (scaled)
1	1	0	0	–1	–1	7, 8
2	0	1	0	–1	–1	4, 4
3	0	0	1	–1	–1	5, 7
4	0.5	0.5	0	–1	–1	7, 8
5	0.5	0	0.5	–1	–1	8, 10
6	0	0.5	0.5	–1	–1	4, 3
7	1	0	0	–1	+1	10, 13
8	0	1	0	–1	+1	8, 8
9	0	0	1	–1	+1	3, 7
10	0.5	0.5	0	–1	+1	12, 16
11	0.5	0	0.5	–1	+1	9, 13
12	0	0.5	0.5	–1	+1	7, 10
13	1	0	0	+1	–1	10, 12
14	0	1	0	+1	–1	5, 8
15	0	0	1	+1	–1	9, 8
16	0.5	0.5	0	+1	–1	14, 15
17	0.5	0	0.5	+1	–1	12, 11
18	0	0.5	0.5	+1	–1	8, 7
19	1	0	0	+1	+1	6, 5
20	0	1	0	+1	+1	7, 4
21	0	0	1	+1	+1	6, 7
22	0.5	0.5	0	+1	+1	5, 5
23	0.5	0	0.5	+1	+1	4, 6
24	0	0.5	0.5	+1	+1	7, 8

[a](Actual run order randomized)
[b](Each combination replicated on random basis.)