Mixing it up with Computer-Aided Design

by Mark J. Anderson, Patrick J. Whitcomb, Stat-Ease, Inc.

With computer software, formulators can take advantage of a powerful statistical tool: design of experiments (DOE) for mixtures. DOE methods employ test arrays that produce maximum information from minimal runs. Industrial experimenters typically turn to two-level factorials as their first attempt at DOE. These designs consist of all combinations of each factor at its high and low levels. With large numbers of factors, only a fraction of the runs need to be completed to produce estimates of main effects and simple interactions[1]. However, when the response depends on proportions of ingredients, such as in chemical or food formulations, factorial designs may not make sense. For example, look at what happens with experiments on lemonade (Table 1).

Table 1: Misleading Factorial Design for Lemonade

Run	Lemons	Water (cups)	Ratio Lemons/Water	Taste
1	1	1	1.0	Good
2	2	1	2.0	Sour
3	1	2	0.5	Weak
4	2	2	1.0	Good

Runs 1 and 4 do not provide any taste contrast. It would make more sense to look at taste as a function of the ratio of lemons to water. Mixture design accounts for the dependence of response on proportionality of ingredients. If you formulate chemicals, food or other products, consider using mixture design rather than factorials or related optimization methods. To show you how, follow along as we conduct a kitchen chemistry experiment on pound cake.

Set Up the Experiment

To design the “perfect” pound cake we first chose experimental ranges for each of the four basic ingredients (Table 2).

Table 2: Recipe for Pound Cake Experiment

Component	Ingredient	Range in Ounces
A	Cake Flour[a]	0-5*
B	All-purpose Flour[b]	0-5*
C	Sugar	3-5
D	Butter	3-5
E	Eggs	3-5

[a] Softasilk® from General Mills, Incorporated, Minneapolis
[b] Gold Medal from General Mills, Incorporated, Minneapolis
*Total flour constrained from 3 to 5 ounces ($3 \le (A + B) \le 5$)

Each cake weighed 16 ounces (one pound). The mixture design incorporated two varieties of flour. Cake flour, recommended by experts, costs about four times more than the all-purpose variety. By including both, subject to a multicomponent constraint (see starred footnote to Table 2), we opened the door to a cheaper recipe. Based on a suggestion from a cookbook[2], we added a constant amount of milk (for moisture) and baking powder (for tenderness). These ingredients never changed, so they did not factor in the study.

We entered the components and constraints into a commercially available software package[3]. It provided an optimal set of experiments to reveal interactions between ingredients. We hoped to find synergistic combinations that would lead to a breakthrough pound cake. Cornell[4] provides a very thorough treatment of the design and analysis of mixture designs. Workshops also provide an avenue for developing an in-depth understanding of statistical tools,[5] such as optimal design.

We used four-cavity, no-stick, mini-loaf pans for baking (see Appendix 1 for procedure). Only four of these pans, with a total of sixteen cavities, would fit in our oven, so we split the runs into two blocks. Block one included the extremes plus several repeats of the overall midpoint (centroid) blend. Block two included the midpoints of the binary extremes plus several more repeats of the midpoint blends to tie the two blocks together. Within each block, the software randomized the run order. Randomization provides insurance against lurking factors such as aging ingredients, increasing ambient temperature, and the like. Appendix 2 lists the actual experiment design, which called for 32 blends.

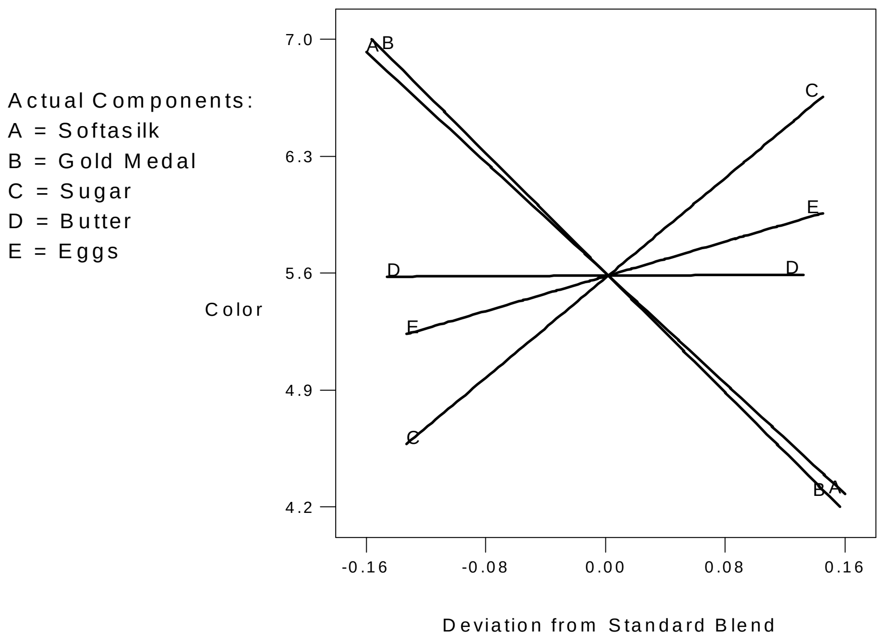
A commercial frozen pound cake provided a standard for measurement of three responses: color, density and taste.

Perform a Statistical Analysis

The statistical results from regression analysis of each response can be seen in Appendix 3. We discovered significant linear fits in every response, with no significant lack of fit.

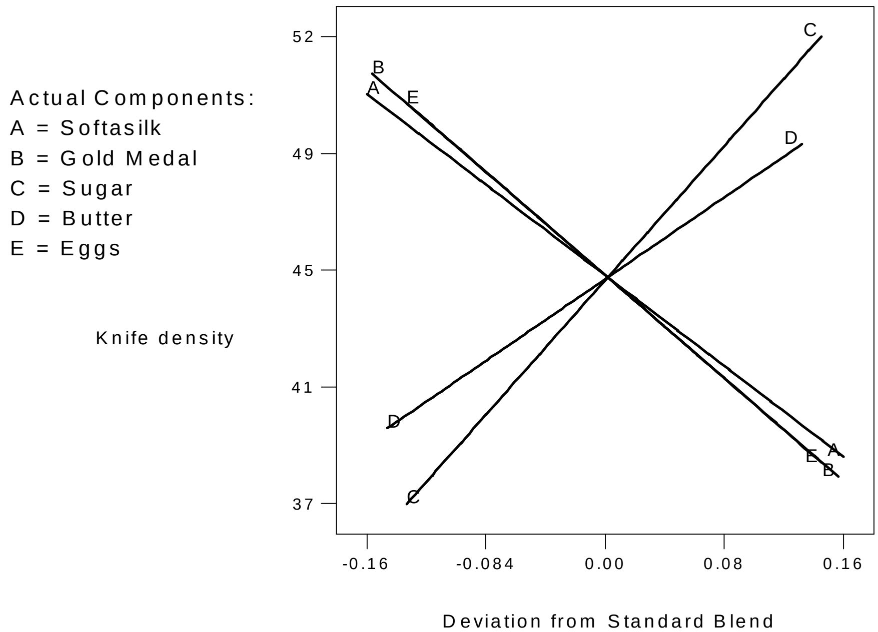
We found no significant interactions. Figures 1, 2 and 3 show trace plots for each of the three responses. Trace plots provide an effective display of the effects of each component on the response. They are especially useful in cases like this, where linear models proved adequate. The plots are constructed by changing the proportion of one component while holding constant the relative proportions of the other components.[4,5]

Figure 1: Trace Plot of Color



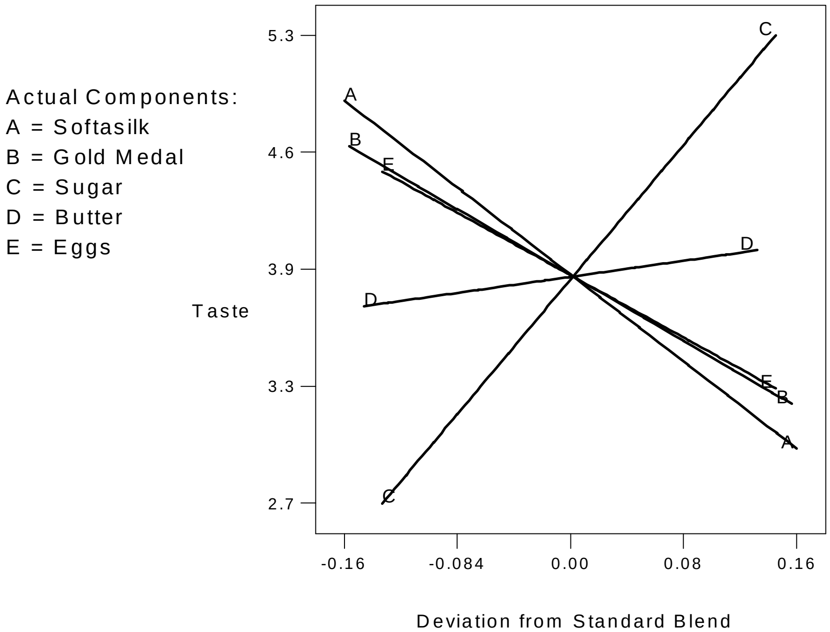
The data on color shows that as flour level increases, the cake surface becomes lighter. All other ingredients cause a darkening in color. The commercial pound cake received a rating of 6.

Figure 2: Trace Plot of Density



The density data shows that increasing flour and eggs decreases the knife penetration, i.e., the cake becomes more dense. Conversely, the addition of sugar and butter increases knife penetration. The commercial pound cake produced a knife density result of 35. None of the experimental cakes approached this level.

Figure 3: Trace Plot of Taste

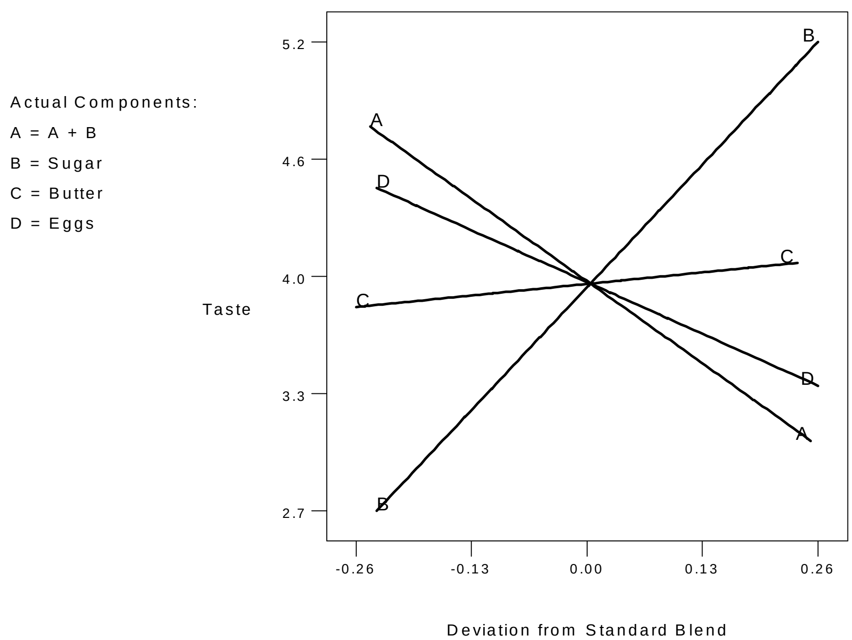


In regard to taste, the favorite ingredient proved to be sugar. Butter might also be preferred, but more tests would be needed to prove this conclusively. High levels of flour met with disapproval. Eggs also may be negatively correlated with taste, but the statistics were not conclusive. The commercial pound cake was rated 7 for taste. We hoped to produce the best taste rating possible, subject to meeting minimum requirements on color and density.

Achieve Big Savings in Raw Material

Regarding type of flour, we saw no difference in any of the responses. Therefore we decided to use all-purpose flour, rather than the expensive cake flour, in our pound cake recipe. For purposes of further analysis, with the aid of software, we combined the two types of flour in to one component. This simplifies presentation of the results, such as the outcome for taste (Figure 4).

Figure 4: Trace Plot of Taste (Combined Flour Components



The use of multicomponent constraints in this manner offers formulators tremendous opportunities for cost reduction. You can use highly efficient mixture designs to simultaneously optimize your component levels while identifying cheaper substitute materials.

Putting It All Together

With statistically valid predictive models in hand for all the key responses, it now became possible to do a multiple response optimization. We performed a desirability analysis[6] to establish the optimal recipe using software designed for this purpose. You can see our specifications in Table 3.

Table 3: Desirability Specifications

Response	Goal	Low	High	Importance	Result
Color	6	4	8	Lowest (+)	6.0
Density (penetration)	Minimize	35	70	Medium(++)	46.0
Taste	Maximize	0	10	Normal (+++)	4.6

We gave taste the most importance as a response, followed by density and then color.

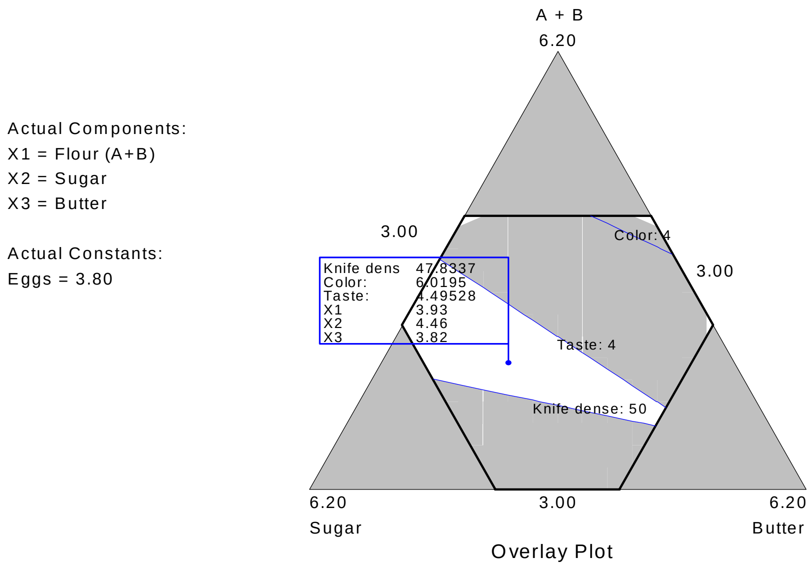
Given these inputs, the software set up desirability scales ranging from zero to one for each response. For example, the taste result of 4.6 got a desirability rating of 0.46 (4.6 out of 10). The individual desirabilities for color and density were calculated in a similar way, but downgraded with lower importance settings.[7] Then the program searched for a recipe that would produce the highest possible overall desirability, computed by multiplying the individual desirabilities and taking the cube root. Color matching proved to be no problem, but the results for density and taste fell short of the ideal set by the commercial product. The optimal recipe found by the program is shown in Table 4.

Table 4: The Most Desirable Recipe for Pound Cake

Component	Ingredient	Ounces
A	Cake Flour	0.0
B	All-purpose Flour	4.3
C	Sugar	4.9
D	Butter	3.0
E	Eggs	3.8

In order to get a picture of the operating window, we then created an overlay response plot with unwanted results shaded (Figure 5). Based on the outcome from the numerical optimization, we tightened up the specifications. The boundaries are taste at a minimum of 4 and knife density at a maximum of 50, with color between 4 and 6. The taste and density define the space. Color did not cause any restrictions. Only three components at a time can be plotted on the triangular plots, so we set the eggs at a fixed level of 3.8 ounces (the optimal level). A flag is set in the middle of the "sweet spot". It shows the predicted response at that formulation.

Figure 4: Overlay Plot Showing Operating Window



The resulting operating window can literally be termed the "sweet spot," since sugar levels press up against the outer boundary. (Note: as you move from the side of the triangle labeled X2 toward the opposing vertex, the level increases from 3 ounces to 9.4 ounces.)

Further work might be done to push the boundary on sugar to see if taste can be improved. Other ingredients might be considered. Eventually we could look at process factors, such as agitation, time, temperature and the like.

Conclusion

A statistical approach to mixture experimentation, aided by computer software, provides major improvements and profound insights within a relatively short time. Formulators who use factorial design or related DOE methods, or worse yet, the old-fashioned one-factor-at-a-time approach, will profit by using mixture design methods.

Acknowledgments:

Thank you to Emily Anderson who baked the cakes. Also, thanks to Louise Jacques of Robin Hood Flour, Jerry Kessler of Quaker Oats, and Jasbir Singh of Sara Lee who provided technical advice.

References:

1. Anderson, Mark, *DOE as Easy as Popping Corn*, PI Quality, July/August 1993, pp 30-32.

2. Beranbaum, Rose Levy, *The Cake Bible*, William Morrow and Company, New York, 1988, pages 23-26.
3. Design-Expert is a trademark of Stat-Ease Corporation, 2021 East Hennepin, Suite 191, Minneapolis, MN 55413. (This software currently runs under Microsoft Windows. The cost is $995.)
4. Cornell, John, *Experiments with Mixtures*, 2nd Edition, John Wiley, New York, 1990.
5. Smith, Wendell, *Mixture Experiments Hold the Keys to Formulation*, Today's Chemist at Work, February 1996, pp 18-24.
6. Mark J. Anderson, Patrick J. Whitcomb, "Optimizing Formulation Performance with Desirability Functions" paper presented at Canadian Metallurgists Conference, Quebec City, August, 1992.
7. Helseth, et al, *Design-Expert Software User's Guide*, Stat-Ease Corporation.

Appendix 1: Procedure for Making Pound Cake

1. Blend the dry ingredients (flour and sugar).
2. Add butter (presoftened in a microwave) and eggs (premixed).
3. Beat the whole blend for 1 minute to develop structure.
4. Place the blended raw materials in a loaf pan.
5. Bake for 50 minutes at 350 degrees Fahrenheit.

Appendix 2: Pound Cake Mixture Design

#	Block	Cake Flour Oz.	Reg Flour Oz.	Sugar Oz.	Butter Oz.	Eggs Oz.	Density[a] mm Penetrate	Color[b] Rating	Taste[c] Rating
1	1	5.00	0.00	3.00	5.00	3.00	23	5	1
2	1	0.00	3.00	5.00	3.00	5.00	30	6	6
3	1	3.00	0.00	3.00	5.00	5.00	34	8	2
4	1	2.50	2.50	3.00	3.00	5.00	35	3	3
5	1	2.00	2.00	4.00	4.00	4.00	43	6	3
6	1	0.00	5.00	3.00	3.00	5.00	22	5	3
7	1	0.00	3.00	5.00	5.00	3.00	58	8	8
8	1	2.00	2.00	4.00	4.00	4.00	39	7	7
9	1	0.00	4.00	3.00	5.00	4.00	36	5	4
10	1	0.00	4.00	3.00	5.00	4.00	44	5	5
11	1	1.50	1.50	3.00	5.00	5.00	61	6	5
12	1	2.00	2.00	4.00	4.00	4.00	35	6	3
13	1	0.00	5.00	3.00	5.00	3.00	31	4	3
14	1	5.00	0.00	5.00	3.00	3.00	61	6	5
15	1	2.50	2.50	5.00	3.00	3.00	41	6	5
16	1	5.00	0.00	3.00	3.00	5.00	16	4	2
17	2	4.00	0.00	4.00	5.00	3.00	55	5	3
18	2	2.00	2.00	4.00	4.00	4.00	68	4	2
19	2	4.00	0.00	4.00	5.00	3.00	76	3	5
20	2	0.00	4.00	5.00	4.00	3.00	65	5	4
21	2	0.00	3.00	4.00	4.00	5.00	46	7	3
22	2	0.00	3.00	4.00	4.00	5.00	52	7	3
23	2	0.00	5.00	4.00	3.00	4.00	34	6	2
24	2	3.00	0.00	5.00	4.00	4.00	46	8	4.5
25	2	2.00	2.00	4.00	4.00	4.00	34	5	4.5
26	2	3.00	0.00	5.00	4.00	4.00	64	8	6
27	2	2.00	2.00	4.00	4.00	4.00	33	7	4
28	2	1.50	1.50	5.00	3.00	5.00	60	8	4
29	2	2.50	2.50	3.00	5.00	3.00	50	3	3.5
30	2	0.00	5.00	4.00	3.00	4.00	45	3	4
31	2	2.00	2.00	4.00	4.00	4.00	35	4	3
32	2	1.50	1.50	5.00	5.00	3.00	63	6	5

[a] Density measured by the penetration in millimeters of a butter knife dropped from height of 12 inches from tip of knife to top of loaf.
[b] Color rated subjectively from 1 (lightest) to 10 (darkest)
[c] Taste rated subjectively on a scale of 1 (worst) to 10 (best) by author.

Appendix 3: Table of Adjusted Effects*
(Probabilities shown in parentheses)

	Color	Density	Taste
Model	Linear (< 0.001)	Linear (0.014)	Linear (0.001)
A	-5.2 (< 0.001)	-21.7 (0.055)	-3.8 (0.002)
B	-5.5 (< 0.001)	-24.7 (0.028)	2.6 (0.026)
C	2.5 (< 0.001)	16.2 (0.009)	2.8 (<0.001)
D	0.5 (0.290)	10.7 (0.053)	0.6 (0.277)
E	1.2(0.026)	-8.3 (0.139)	-0.8 (0.172)

*See Cornell page 254 for details on calculation of adjusted effects.