

©2004 – RICHARD BURNHAM

Design of Experiments Training Banishes Purple-Paper Blues

by
Richard Burnham

Because paper mill trials are extremely expensive, statistical experiments are essential for pre-defining appearances and shades of paper before full production. It's the best way to avoid costly surprises – like purple paper instead of white. So in a Stora Enso North America classroom located close to the company's Wisconsin research facility, company scientist and trainer Donald R. Smith begins another coaching session in how to prevent purple paper – and faces. The training scenario here is somewhat lightheartedly adapted from a real event "to protect the innocent and entertain the students," Smith says with a grin. But the training message is clear – an unexpected interaction takes place when a vendor's new blue dye is added to a coating formulation containing a starch binder, as opposed to a latex binder. Binders help coatings adhere to paper, and the starch binder's unexpected interaction with the new blue dye is very evident in the purple paper confronting the students. The interaction's discovery vehicle is design of experiments (DOE), and the primary lesson Smith is driving home in this training segment is the power of DOE over traditional one-factor-at-a-time (OFAT) testing.

"In a typical mill trial, even one that is designed to impact appearance, we encourage practitioners to evaluate any responses that could possibly be affected," Smith recently shared. With this in mind, let's look at a portion of a recent DOE training session Smith regularly skips (see sidebar).

FIVE RESPONSES, FOUR FACTORS

Although a paper's optical properties can be defined by several possible responses (the outcomes from testing varying factors), this training DOE concentrates on five particularly valuable color-evaluation responses:

1. Brightness (reflective intensity)
2. a-Color (red and green components)
3. b-Color (yellow and blue components)

4. L-Value (lightness intensity)
5. CIE Whiteness (combines a-Color, b-Color, and L-Value to represent visual appearance better than brightness can, although it doesn't communicate true color or reflectivity).

In the training scenario, two engineers, both following proper quality control procedures with statistically valid results on the same dye, produced two conflicting mill trial outcomes. The first engineer showed that the new dye's impact on a-Color and other optical properties was negligible. However, the second engineer found that the same dye so dramatically increased a-Color, that it could not be corrected by adjusting the red-dye level. The paper turned purple. So did the production manager, according to legend.

Stora Enso's Quality Systems engineers, knowledgeable in experimental design, examined the data from both engineers. They discovered that different reasons for conducting the two trials had led to different binder choices. Although both studied the same blue dye, one engineer used starch as a binder – the other used latex. The results from both mill trials were valid, but neither trial could see hidden interactions because the factors were studied one at a time. Factor levels must be changed in combination to reveal interactions – something DOE readily exposes.

Design of experiments methods are used to analyze many variables with ease by breaking their study into multiple experimental designs. Because this is a training DOE leading to more complex scenarios, just four factors are under study in this DOE, called a full factorial (Table 1).

FACTOR	UNIT	LOW LEVEL	HIGH LEVEL
Delaminated Clay	%	0	100
Latex Binder	%	0 (0% latex, 100% starch)	100 (100% latex, 0% starch)
Titanium dioxide (TiO ₂)	Parts	0	10
Blue Dye	ml	0	1000

TABLE 1. FACTORS AND LEVELS – The levels are set beyond the extremes of their normal parameters to exercise the statistical power of DOE. Stora Enso researcher and trainer Donald Smith says to students, “We want you to push beyond what you think is normally achievable. We encourage this primarily because we don't want you to lock into a pre-defined box.”

A primary benefit with full-factorial experimentation is that all possible combinations of all factors are studied at their designated levels. This differs from subsequent, more complex, classroom exercises called fractional factorials that study factors and subsets of combined factor levels. Fractional-factorial DOE's can cause some factor effects to be confounded, or aliased, with other factors. They are often used for screening dozens of factors to find the significant few that are key. Still easily solvable with sophisticated DOE software, fractional factorials first require an understanding of full factorials.

Testing each of the four factors at two levels yields 16 experiments ($2 \times 2 \times 2 \times 2 = 16$). Figure 1 shows the DOE design under evaluation in the training session as it appears in a commercially available DOE product, Design-Expert® software. The column marked "Std" indicates conventional array ordering. "Run" lists the experiments as originally performed, in randomized order, to counteract any lurking variables such as ambient humidity. "Block 1" in this situation means all experiments were conducted using resources not affected by differing days, methods, or materials. A seventeenth experiment was conducted, using the "center points" of the four factors. (Center points test for curvature by measuring the departure of a response surface from a flat surface – a topic Smith explores in the classroom but is beyond the scope of this article. In short, curvature statistically confirms the DOE's validity.)

	Std	Run	Block	Factor 1 A:A: % Delam %	Factor 2 B:B: % Latex %	Factor 3 C:C: TiO2 parts	Factor 4 D:D: Blue Dye ml	Response 1 Brightness	Response 2 a Color	Response 3 b Color	Response 4 L Value	Response 5 CIE Whiteness
	17	1	Block 1	50.00	50.00	5.00	500.00	81.25	0.78	0.96	90.76	76.5767
	11	2	Block 1	0.00	100.00	0.00	1000.00	76.13	0.82	0.91	87.78	72.2433
	1	3	Block 1	0.00	0.00	0.00	0.00	78.08	0.23	4.93	91.84	59.7633
	14	4	Block 1	100.00	0.00	10.00	1000.00	82.82	3.28	-2.05	89.56	88.71
	5	5	Block 1	0.00	0.00	10.00	0.00	82.16	0.16	4.12	93.58	67.0467
	2	6	Block 1	100.00	0.00	0.00	0.00	81.87	0.27	3.64	93.05	68.4867
	9	7	Block 1	0.00	0.00	0.00	1000.00	74.99	3.6	-0.86	85.86	76.6767
	7	8	Block 1	0.00	100.00	10.00	0.00	81.62	0.08	4.44	93.51	66.2067
	6	9	Block 1	100.00	0.00	10.00	0.00	85.98	0.23	2.82	94.75	75.8167
	10	10	Block 1	100.00	0.00	0.00	1000.00	78.7	3.47	-1.9	87.3	84.11
	3	11	Block 1	0.00	100.00	0.00	0.00	77.62	0.24	5.16	91.74	59.14
	4	12	Block 1	100.00	100.00	0.00	0.00	81.64	0.39	3.64	92.92	68.84
	15	13	Block 1	0.00	100.00	10.00	1000.00	79.76	0.75	0.7	89.75	75.9867
	16	14	Block 1	100.00	100.00	10.00	1000.00	83.82	0.97	-0.73	91.01	84.7833
	13	15	Block 1	0.00	0.00	10.00	1000.00	79.81	3.34	-1.02	88.57	81.97
	8	16	Block 1	100.00	100.00	10.00	0.00	85.41	0.23	2.89	94.5	75.1667
	12	17	Block 1	100.00	100.00	0.00	1000.00	79.88	1.07	-0.52	88.92	80.87

FIGURE 1. FACTORS AND RESPONSES – Screenshot displays the DOE's design matrix, listing 17 experiments conducted with four factors and five responses under study in this classroom DOE. The 17th standard run in the first column marked "Std" is a simplified "center point" experiment to introduce the concept to students. For ideal DOE analysis, four center point runs are best – and

subsequently studied in training. The software's menu at left allows users to analyze, and then optimize, any of the five responses. (Figure screenshot courtesy of Stat-Ease, Inc.)

INTERACTION ANALYSIS

The full-factorial DOE reveals that each of the five responses has at least one significant interaction. But the most critical interaction is the one that shows up in the a-Color response. Factor levels producing optimal brightness and whiteness interact differently with starch versus latex, regrettably affecting a starch binder's a-Color to the point that it appears as purple – not a good result for a customer who wants bright white paper.

"It's well known in our field that to increase brightness, you increase the blue color, but going too far can produce purple paper. The good news is that both engineers were right," says Smith, "but they both missed the [latex factor] interaction. That drives home the value of doing a designed experiment." Figure 2 describes and shows how one-factor-at-a-time methods produced two opposing mill trials.

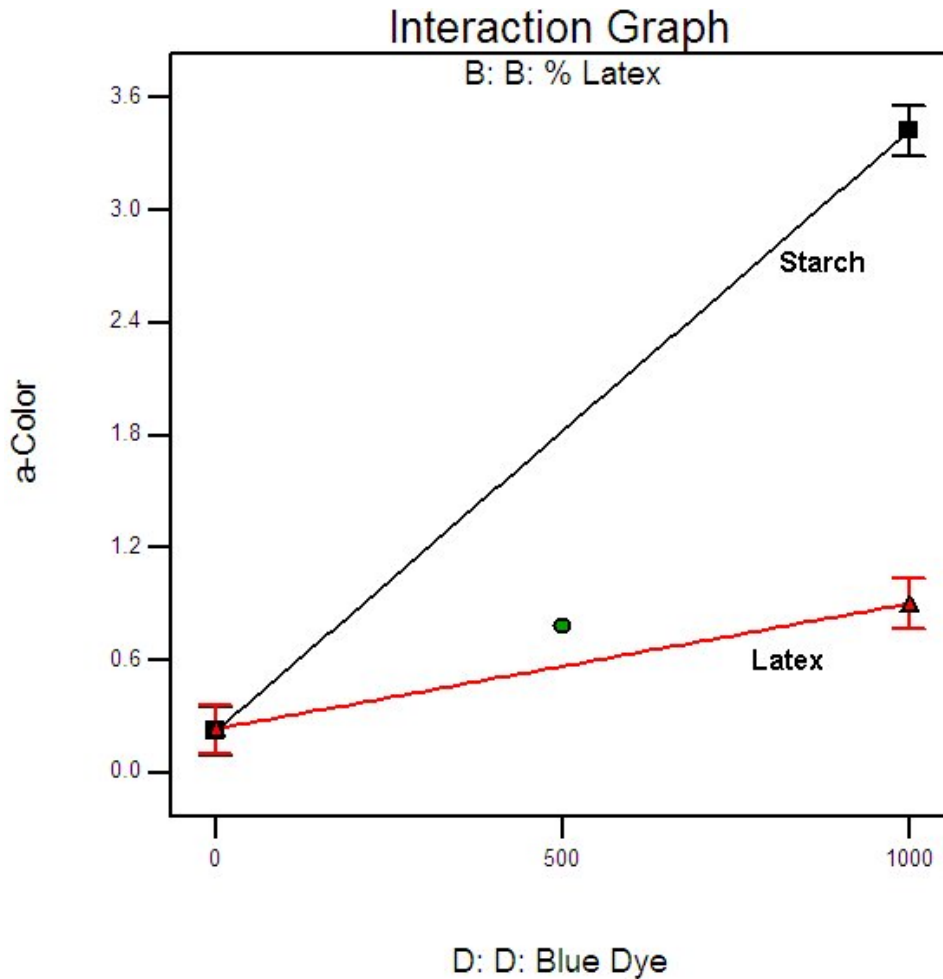


FIGURE 2. INTERACTION GRAPH – Interactions occur when the outcome of a factor’s response differs depending on the setting of another factor being tested. Here, the two lines are not parallel, confirming that an interaction is occurring. The red line (a-Color response) shows that in the presence of the new blue dye, as latex binder decreases from 1000 ml (0% starch binder) to 0 ml latex (100% starch), a-Color values soar, an interaction that turns white paper purple. It’s an important lesson showing how OFAT blocks breakthrough discoveries. (Figure screenshot courtesy of Stat-Ease, Inc.)

The interaction was misinterpreted as a conflict between the two engineers’ OFAT data. As the training scenario comes to an end, the story goes that both engineers were considerably relieved that only a new mill trial was chosen for execution.

“CLASSROOM TO LAB TO MILL FOR DOE-OPTIMIZED PAPER COATINGS”

Based on real-life paper-coating studies at Stora Enso's Biron Research Centre in Wisconsin Rapids, Wisc., scientists regularly conduct classroom training in design of experiments (DOE). Their objective is to provide engineers, researchers, and technicians with the skills to design and analyze DOEs that overcome the inherent weakness of one-factor-at-a-time (OFAT) experimentation. OFAT doesn't reveal factor interactions. Stora Enso quality engineer and DOE trainer Donald R. Smith says, "There is significant, bottom-line value in pursuing better trials. DOE is an invaluable tool for survival and success."

DOE uses highly efficient mathematical processing schemes in tandem with experimental matrices. It reveals how a total system works by providing information not only about individual factors, but also about their interactions. And given the high cost of experimentation, DOE's efficiency is reason enough to abandon OFAT testing. Says Smith, "When we're performing a DOE, almost exclusively what we're trying to discern is what the effect is on output variables when we vary input component factors. We want to be able to predict the brightness, for example, based on whatever the amount is of each factor."

The accompanying story follows Smith during a portion of the DOE class he teaches. It illustrates how two conflicting OFAT results are resolved when conducting a standard, two-level, full-factorial DOE. The analysis reveals a previously undetectable factor interaction that's invisible using OFAT methods.

ABOUT THE AUTHOR:

The author would like to acknowledge the following for their help with this article: Stora Enso's Donald R. Smith and Carol Davis, and Stat-Ease's Mark Anderson. Richard Burnham has written, co-authored, or ghostwritten over 285 technical articles about companies and individuals. He is the founder of Publication Coordination, a writing firm specializing in trade publication articles. He may be reached at bigbooker@gmail.com.

CONTACT INFORMATION:

Richard Burnham

2700 Henry St. N.

St. Paul, MN 55109-1858

Phone: (920) 699-2904

E-mail: bigbooker@gmail.com

Mr. Donald R. Smith

Stora Enso North America

Research Scientist, Quality Systems

Biron Research Centre

PO Box 8050

Wisconsin Rapids, WI 54495

Phone: 715-422-2456 Fax: 715-422-2458

E-mail: don.smith@storaenso.com

Web site: www.storaenso.com

Stat-Ease, Inc.

Design-Expert® software

2021 East Hennepin Avenue, Suite 480

Minneapolis, MN 55413

Phone: (612)378-9449 Fax: (612)378-2152

E-mail: info@statease.com

Web site: www.statease.com
