This item is the archived peer-reviewed author-version of:

A comparison of two-level designs to estimate all main effects and two-factor interactions

Reference:

Bradley Jones, Schoen Eric D., Montgomery Douglas C..- A comparison of two-level designs to estimate all main effects and two-factor interactions

Quality engineering - ISSN 0898-2112 - 28:4(2016), p. 369-380

Full text (Publishers DOI): <http://dx.doi.org/doi:10.1080/08982112.2015.1136829>

To cite this reference: <http://hdl.handle.net/10067/1357600151162165141>

A Comparison of Two-Level Designs to Estimate all Main Effects and Two-Factor Interactions

Bradley Jones^{1,2}, Eric D. Schoen^{2,3} and Douglas C. Montgomery⁴

¹SAS Institute, Cary NC, USA

²University of Antwerp, Belgium

³TNO, Zeist, Netherlands

⁴Arizona State University, Tempe AZ, USA

Abstract

We compare cost efficient alternatives for the full factorial 2^4 design, the regular 2^{5-1} fractional factorial design and the regular 2^{6-1} fractional factorial design that can fit the model consisting of all the main effects as well as all the two-factor interactions. For 4 and 5 factors we examine orthogonal arrays with 12 and 20 runs, respectively. For 6 factors we consider orthogonal arrays with 24 as well as 28 runs. We consult complete catalogs of two-level orthogonal arrays to find the ones that provide the most efficient estimation of all the effects in the model. We compare these arrays with D-optimal designs found using a coordinate exchange algorithm. The D-optimal designs are always preferable to the most efficient orthogonal arrays for fitting the full model in all the factors.

KEY WORDS: A-optimal design; D-optimal design; Orthogonal Array; Strength 2; [Column Balance](#)

1. Introduction

This paper considers the design of experiments for four to six two-level factors to fit a model consisting of the main effects and all the two-factor interactions (2FIs). We call this model the interactions model. Well-known candidate designs are the full factorial 2^4 design in 16 runs, the regular 2^{5-1} fractional factorial design in 16 runs and the regular 2^{6-1} fractional factorial design in 32 runs. The design for five factors can be constructed by writing down a full factorial design in four factors coded with -1 and 1, or – and +, and calculating the settings of the fifth factor by multiplication of the settings of the first four factors. [For example, if the settings of the first four](#)

factors are (+ - + +), the setting of the last factors is $1 \cdot -1 \cdot 1 \cdot 1 = -1$. The fractional design for six factors can be constructed similarly.

There are several reasons to like the above candidate designs. It is easy to calculate the effect of changing any factor: all the responses for trials run at the $-$ level of the factor are averaged and subtracted from the average of the responses for trials run at the $+$ level of the factor. The interaction effects are also easy to calculate: the settings of the two factors involved in an interaction are multiplied to obtain a new column with $+$ and $-$ levels, and the calculation of interaction effects is analogous to the calculation of main effects. Most importantly, the effects of any factor or interaction can be estimated with the smallest possible variance of all designs having the same number of runs (i.e. σ^2/n where n is the number of runs and σ^2 is the variance of the error term in the model). These effects achieve this lower bound because they are mutually orthogonal making these designs both D- and A-optimal. The column for each effect is also balanced having an equal number of plusses and minuses. If two effect columns are correlated, the variance of each effect must be greater than σ^2/n . In addition, correlated effect columns require a regression analysis to calculate the estimates of the main effects and 2FIs.

In the four-factor case, there are 11 unknown quantities that we want to estimate (i.e. the intercept, four main effects, and six two-factor interactions). Mee (2004) calls the ratio of the number of unknowns to the number of runs the degree of freedom efficiency. For the full-factorial design this efficiency is 11/16. However, any 4-factor-column subset of the Plackett-Burman design having 11 factors and 12 runs (Plackett and Burman, 1946) can also fit the interactions model though not with minimum possible variance for each parameter. Such a design, which we call the PB4 design, has a degree of freedom efficiency of 11/12.

The PB4 design is an example of an orthogonal array of strength 2. A strength-2 orthogonal array (OA) is composed of a matrix of symbols, in our case $+$ and $-$, such that any group of two columns

has an equal number of all possible combinations of the symbols. Specifically, any pair of columns in the PB4 design has 3 replications of the four possible pairs of symbols: + +, + -, - +, and - -.

Hedayat et. al. (1999) provide a comprehensive treatment of OAs and their properties. In our context the most important property is that an OA of strength 2 allows for fitting all the main effects with minimum variance. ~~However, w~~While the PB4 design also allows for fitting the interactions model, this is not generally true for all strength 2 OAs. Moreover, even if a strength 2 OA can fit a model containing all these interactions, the resulting effects ~~will~~may be correlated and therefore, their variances ~~will~~may no longer be minimal ~~(i.e. σ^2/n where n is the run budget).~~

In the five factor case there are 16 unknown quantities that we want to estimate (the intercept, five main effects and ten 2FIs). So the degree of freedom efficiency for the 2^{5-1} design is 100%. However, there are no degrees of freedom left to estimate the error variance. Therefore, it might ~~makes~~ sense to consider designs with additional runs to provide the extra degrees of freedom needed for statistical inference. The most obvious candidates would be designs based on orthogonal arrays of 20 runs. Such designs have a degrees of freedom efficiency of 16/20 or 0.8, but it is necessary to check whether they can fit the interactions model.

For the six factor case, the degree of freedom efficiency for the 2^{6-1} design is only 22/32. When runs are costly, it makes sense to investigate the properties of alternative designs having fewer than 32 runs. There are surprisingly many strength-2 OAs for six two-level factors having either 24 or 28 runs, and we are interested in those that best fit the interactions model.

In this article, we consult a catalog of all the different strength-2 OAs for four factors in 12 runs, five factors in 20 runs, and six factors in 24 and 28 runs (Schoen, Eendebak, and Nguyen, 2010) to find the best ones to use as a design for fitting the interactions model. Recall that these OAs minimize the variances ~~and correlations~~ among the main effects only. However, our concern is with their performance for estimating the two-factor interactions.

Since we are interested in finding a design that minimizes the variances of all the model coefficients, we also employ a ~~D-optimal~~ design generation algorithm that aims for generally to lower the variances of the effects ~~in a global sense~~. We then compare the resulting designs with the best OAs.

When comparing designs, it is important to be clear about the design features that are desirable.

Orthogonal arrays were first introduced and were desirable because in the absence of computers, it is possible to calculate all the effects using differences of averages of the responses. Additional desirable features of OAs are their orthogonal main effects and balanced columns. Balanced and orthogonal columns are useful design properties allowing each effect to be evaluated separately without consideration of any other effect. Because of this there is a commonly held belief that orthogonality provides model robustness. This is true but only for the set of effects that are orthogonal. For strength-2 OAs, the 2FIs are not all orthogonal, so it may not be possible to estimate the interactions model. Even when the interactions model is estimable the variances of the estimates of the 2FIs may be large.

In this work we contrast OAs with computer-generated designs constructed to fit the interactions model. For many of these computer-generated designs the main effects may not be orthogonal and the columns of the design may not be balanced. In such cases these designs do not estimate the main effects with full efficiency. However, in exchange for the loss of precision for estimating main effects, these designs can often provide more precise estimates of the 2FIs. When there is concern that there may be many active 2FIs, then it may be worthwhile to trade the loss of precision in the estimates of the main effects for more precise estimates of the 2FIs.

The organization of the article is as follows. In Section 2, we review D-optimality and A-optimality to quantify the capacity of a design to give precise estimates of the model parameters. In Section 3, we discuss the results of our search for optimal designs. We judge the optimality of the design alternatives with respect to the interactions model. In Section 4, we compare ~~the various~~the efficiency and model robustness of six-factor designs when the number of active factors varies from

two to five, considering two model spaces from the literature. We conclude with a brief discussion and some recommendations.

2. Criteria to Measure Design Efficiency

We make use of two well-known design optimality criteria. A design is said to be optimum with respect to the D-optimality criterion if $|\mathbf{X}'\mathbf{X}|$ is maximized, where $\mathbf{X}$ is the model matrix for the design. That is, $\mathbf{X}$ is the design matrix whose columns have been expanded to model form. For example, the interactions model for a four-factor design requires an $\mathbf{X}$ matrix with a column of ones for the intercept, the four main effect columns and six columns for 2FIs. The D-efficiency of a design is

$$D_{eff} = \left(\frac{|\mathbf{X}'\mathbf{X}|}{\text{Max}|\mathbf{X}'\mathbf{X}|} \right)^{1/p}$$

where the maximization in the denominator is taken over all designs and p is the number of parameters in the model. In cases where the globally D-optimal design is not known, as in all our examples, we replace the expression in the denominator above with n^p . An $\mathbf{X}$ matrix with p orthogonal columns of plus and minus ones would have a determinant of n^p . This makes the result of our replacement a lower bound on the D-efficiency – see Goos and Jones (2011). Generally this lower bound is not tight and the true D-efficiency may be substantially higher than the reported lower bound.

To compare the relative D-efficiency of design 1 to design 2 with respect to a common model, one computes $(|\mathbf{X}_1'\mathbf{X}_1| / |\mathbf{X}_2'\mathbf{X}_2|)^{(1/p)}$. It is possible for design 2 to be less efficient than design 1 even if design 1 has fewer runs.

Fractional-factorial designs and Plackett-Burman designs are known to be both D-optimal and A-optimal (i.e. efficiencies of 100%) for the main effects model. For an arbitrary number of runs, however, the D-optimal design is generally unknown and finding a highly D-efficient design requires

the use of a computer. Exchange algorithms such as the coordinate exchange algorithm of Meyer and Nachtsheim (1995) are widely used to construct D-optimal designs quickly and efficiently.

A design is A-optimal if it minimizes the sum of all of the diagonal elements of $(\mathbf{X}'\mathbf{X})^{-1}$. The diagonal elements of $(\mathbf{X}'\mathbf{X})^{-1}$ are proportional to the variances of the model regression coefficients where the constant of proportionality is σ^2 . Therefore, the A-optimal design minimizes the average variance of the estimates of the parameters. The A-efficiency of a design is

$$A_{eff} = \left(\frac{p}{n \text{Tr}(\mathbf{X}'\mathbf{X})^{-1}} \right)$$

where the $\text{Tr}(Z)$, read trace of Z , is the sum of the diagonal elements of Z , and p is again the number of parameters in the model.

D-optimal designs and D-efficiency as a measure are more familiar terms. However, The A-efficiency is a more direct measure than the D-efficiency of the variances of the model coefficients. For a more in-depth treatment see Atkinson et. al. (2007). The previously mentioned coordinate exchange algorithm is easily adapted to compute designs with high A-efficiency.

3. Efficient Designs for Four, Five and Six Factors

In this section we consider efficient designs for four factors in 12 runs, five factors in 20 runs, and six factors in 24 or 28 runs. We identify designs from a complete catalog of OAs that are best in terms of their D - and A -efficiencies [for the interactions model](#). Then we compare these designs to a D -optimal design [constructed for efficient estimation of the interactions model](#).

3.1. Construction and comparison of 12 run designs for four factors

A natural construction method for a four-factor OA with 12 runs is to choose four factors of the Plackett-Burman design for 11 factors and 12 runs. All choices of 4 out of the 11 factors result in isomorphic designs. We say that two designs are isomorphic if it is possible to convert one to the other using a combination of three operations:

- 1) Switching two rows
- 2) Switching two columns
- 3) Permuting the symbols in a column – i.e. changing the +s to –s and vice versa.

So, two arrays are distinct if no combination of the above actions can change one array to the other. Note that none of these operations has any effect on the statistical properties of a design. However, any two nonisomorphic OAs may have substantially different statistical properties, which makes comparing them worthwhile.

It turns out there is only one OA for four factors and 12 runs (Schoen et al., 2010). This design appears in the right panel of Table 1.

A well-known alternative to this 12 run design is the design constructed as a $\frac{3}{4}$ fraction of a full factorial proposed by John (1962). This is not an orthogonal design, because the four level

combinations of the third and fourth factor do not occur equally often. The middle panel of Table 1 shows John's design.

|

Table 1. Three designs for 4 factors and 12 runs

Run	D-optimal	John	OA
1	+ - - +	+ + - -	- - - -
2	+ + - -	+ - - +	+ + + -
3	- + + +	+ + + -	- + - +
4	+ + + -	- - + -	- + + +
5	- + - -	+ + - +	+ + + -
6	+ + + +	- - - -	+ - - -
7	+ - + +	- + - +	+ - + +
8	+ + - +	- + + +	- + - -
9	+ - - -	- - - +	- - + +
10	- - + -	- + + -	+ - - +
11	- - - +	+ - + +	+ + - +
12	- + - +	+ - + -	- - + -

We created a third option by using the coordinate exchange algorithm of Meyer and Nachtsheim (1995) as implemented in JMP 11 (2013). This option, which we label 'D-optimal' appears in the left panel of Table 1. Like the John design, the D-optimal design is not orthogonal. Moreover, the columns of this design do not have the same number of plusses and minuses. Each of the factors has 7 settings of one kind against 5 settings of the other kind. Interestingly, this option has the same D- and A-efficiencies (85.8% and 69.8% respectively) as the $\frac{3}{4}$ fraction. ~~The D-efficiency for the OA is 81.4%.~~ This example illustrates the fact that there may be multiple designs that are D-optimal for a given model and specified number of runs. Incidentally, The D-efficiency for the OA is 81.4 % and its A-efficiency is 63.8%.

In Table 2, we compare the standard error of model coefficient estimators for the three design options, assuming that the error standard deviation is 1. It is obvious that the OA is inferior to the other two designs for the estimation of the factorial effects. Only the intercept is better estimated by the OA than by the other two designs. We note that for the purpose of screening, the precision of the intercept is of less importance than the precision of the other effects.

Table 2: Relative standard errors for three 12-run 4-factor designs

Parameter	standard error		
	D-optimal	John	OA
Intercept	0.31	0.31	0.29
A	0.34	0.35	0.37
B	0.34	0.35	0.37
C	0.34	0.35	0.37
D	0.34	0.35	0.37
AB	0.35	0.35	0.37
AC	0.35	0.35	0.37
AD	0.35	0.35	0.37
BC	0.35	0.35	0.37
BD	0.35	0.35	0.37
CD	0.35	0.31	0.37

In Table 2, we compare the standard error of model coefficient estimators for the three design options, assuming that the error standard deviation is 1. It is obvious that the OA is inferior to the other two designs for the estimation of the factorial effects. Only the intercept is better estimated by the OA than by the other two designs. We note that for the purpose of screening, the precision of the intercept is of less importance than the precision of the other effects.

For the OA the absolute value of the correlation between each main effect and three 2FIs is one-third. The other 3 possible correlations in each case are zero. So, 12 of the 24 possible absolute correlations are one-third. For the design proposed by John, only 7 of the 24 possible absolute correlations between main effects and 2FIs are one-third. The rest are zero. For the D-optimal design all the main effects are have absolute correlations of 0.169 with all the 2FIs.

Adding the 2FIs to the model means that multiple regression is required to fit the parameters of any of these designs. The length of the confidence intervals on the parameters is directly proportional to the standard errors of the coefficients.

The design proposed by John has equal standard errors (0.35) for all factorial effects except for the CD interaction effect (0.31). This feature is useful only if this effect is known to be of special interest. We slightly prefer the other D-optimal design because all the main effects in this design have lower

standard errors than the John design (0.34 versus 0.35). The interactions have standard errors that are the same as for the John design except for the aforementioned CD interaction effect.

3.2. Construction and comparison of 20 run designs for five factors

A 20-run design in five factors (with 16 effects to estimate) has the benefit of providing four degrees of freedom to estimate the variance of the random error. Schoen et al. (2010) show that there are 11 nonisomorphic five-factor OAs having 20 runs.

It is desirable for a strength-2 OA with an adequate number of runs to fit the interactions model. Investigators employing an OA that cannot fit the interactions model run the risk of biased main effects caused by interactions that cannot be estimated. As a result confirmation runs could fail to validate the fitted model. It turns out that Θ only four of the 11 five-factor arrays can fit the interactions model. We show the best one in terms of D-efficiency -in the right panel of Table 3. Its D- and A-efficiencies for the interactions model are 87% and 73% respectively.

For the OA each main effect has absolute correlations of 0.2 with six of the ten 2FIs. The other four correlations are zero for an average absolute correlation of 0.12.

Using the coordinate exchange algorithm in JMP 11, we created the D-optimal design for 20 runs and five factors shown in the left panel of Table 3. Its D-efficiency for the interactions model is 95.1%. Note that the first and last columns of the D-optimal design are not balanced. This results in absolute values of correlations among the main effect columns that average 0.08. The correlations between main effects and 2FIs for this design are variable with a maximum absolute correlation of 0.2. The average absolute correlation between main effects and 2FIs is 0.065, which is a little over half the comparable quantity for the

We compare the standard error of the estimators of the coefficients in Table 4. With either design, the standard errors for the factorial effects are equal. However, those for the OA are uniformly higher than those for the D-optimal design. For this reason, we prefer the D-optimal design.

Table 3: Design alternatives for five factors and 20 runs for fitting the interactions model		
Run	D-optimal	OA
1	- - - - +	- - - - -
2	- - - + -	- - - - +
3	- - - + +	- - - + -
4	- - + - -	- - + - -
5	- - + + -	- - + + +
6	- - + + +	- + - - +
7	- + - - -	- + - + -
8	- + - + +	- + + - +
9	- + + - -	- + + + -
10	- + + - +	- + + + +
11	- + + + -	+ - - - +
12	+ - - - -	+ - - + +
13	+ - - + +	+ - + - -
14	+ - + - +	+ - + + -
15	+ - + + -	+ - + + +
16	+ + - - -	+ + - - -
17	+ + - - +	+ + - + -
18	+ + - + -	+ + - + +
19	+ + + - -	+ + + - -
20	+ + + + +	+ + + - +

Table 4: Standard errors for two 20-run 5-factor designs

Parameter	standard error	
	D-optimal	OA
Intercept	0.23	0.22
A	0.23	0.26
B	0.23	0.26
C	0.23	0.26
D	0.23	0.26
E	0.23	0.26
AB	0.23	0.26
AC	0.23	0.26
AD	0.23	0.26
AE	0.23	0.26
BC	0.23	0.26
BD	0.23	0.26
BE	0.23	0.26
CD	0.23	0.26
CE	0.23	0.26
DE	0.23	0.26

Adding the 2FIs to the model means that multiple regression is required to fit the parameters of any of these designs. The length of the confidence intervals on the parameters is directly proportional to the standard errors of the coefficients. As a result confidence intervals on every parameter other than the intercept will be 12.7% longer for the OA than for the D-optimal design. This would have an effect on the power to detect a regression coefficient equal in size to the error standard deviation assuming all terms are in the model. For each effect the power for the OA is 0.81 compared to 0.89 for a t-test with a significance level of 0.05.

~~We compare the standard error of the estimators of the coefficients in Table 4. With either design, the standard errors for the factorial effects are equal. However, those for the OA are uniformly higher than those for the D-optimal design. For this reason, we prefer the D-optimal design.~~

3.4.3.3. Construction and comparison of 24 run designs for six factors

A natural and often used construction method for a six-factor OA in 24 runs is to choose six factors from the Plackett-Burman design with 23 factors, which is given in Table 5. Note that

~~The Plackett-Burman design is a strength-2 OA because~~ each of the four pairs of symbols (++, +-, -+ and --) appears six times in every one of the 253 pairs of columns.

Not every choice of six columns from the Plackett-Burman design permits fitting the interactions model. There are 100,947 ways to choose six columns out of the 23. Only 49,588 ways allow for fitting the interactions model. Taking the first six columns gives a design with an average variance more than 7 times bigger than for the best choice of six columns (1-4, 10 and 11). So, constructing these designs by arbitrarily choosing six columns from the Plackett-Burman design is ill--advised.

Table 5: Plackett-Burman design for 24 runs and 23 factors

run	factor																						
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
2	-	+	+	+	+	-	+	-	+	+	-	-	+	+	-	-	+	-	+	-	-	-	-
3	-	-	+	+	+	+	-	+	-	+	+	-	-	+	+	-	-	+	-	+	-	-	-
4	-	-	-	+	+	+	+	-	+	-	+	+	-	-	+	+	-	-	+	-	+	-	-
5	-	-	-	-	+	+	+	+	-	+	-	+	+	-	-	+	+	-	-	+	-	+	-
6	-	-	-	-	-	+	+	+	+	-	+	-	+	+	-	-	+	+	-	-	+	-	+
7	+	-	-	-	-	-	+	+	+	+	-	+	-	+	+	-	-	+	+	-	-	+	-
8	-	+	-	-	-	-	-	+	+	+	+	-	+	+	+	-	-	+	+	-	-	+	+
9	+	-	+	-	-	-	-	-	+	+	+	+	-	+	-	+	+	-	-	+	+	-	-
10	-	+	-	+	-	-	-	-	-	+	+	+	+	-	+	-	+	+	-	-	+	+	-
11	-	-	+	-	+	-	-	-	-	-	+	+	+	+	-	+	-	+	+	-	-	+	+
12	+	-	-	+	-	+	-	-	-	-	-	+	+	+	+	-	+	-	+	+	-	-	+
13	+	+	-	-	+	-	+	-	-	-	-	-	+	+	+	+	-	+	-	+	+	-	-
14	-	+	+	-	-	+	-	+	-	-	-	-	-	+	+	+	+	-	+	-	+	+	-
15	-	-	+	+	-	-	+	-	+	-	-	-	-	-	+	+	+	+	-	+	-	+	+
16	+	-	-	+	+	-	-	+	-	+	-	-	-	-	-	+	+	+	+	-	+	-	+
17	+	+	-	-	+	+	-	-	+	-	+	-	-	-	-	-	+	+	+	+	-	+	-
18	-	+	+	-	-	+	+	-	-	+	-	+	-	-	-	-	-	+	+	+	+	-	+
19	+	-	+	+	-	-	+	+	-	-	+	-	+	-	-	-	-	-	+	+	+	+	-
20	-	+	-	+	+	-	-	+	+	-	-	+	-	+	-	-	-	-	-	+	+	+	+
21	+	-	+	-	+	+	-	-	+	+	-	-	+	-	+	-	-	-	-	-	+	+	+
22	+	+	-	+	-	+	+	-	-	+	+	-	-	+	-	+	-	-	-	-	-	+	+
23	+	+	+	-	+	-	+	+	-	-	+	+	-	-	+	-	+	-	-	-	-	-	+
24	+	+	+	+	-	+	-	+	+	-	-	+	+	-	-	+	-	+	-	-	-	-	-

Schoen et al. (2010) show that there are 1,350 nonisomorphic six-factor OAs having 24 runs. The best D- efficiencies and A-efficiencies of these designs for the interactions model are listed in Table

6, together with those of the best design created using the coordinate exchange algorithm in JMP

~~110 with 10,000 random starts.~~

Since there are only 1,350 OAs in six factors, most of the 100,947 six-column choices from the 23

columns of the Plackett-Burman design in Table 5 ~~are not are isomorphic to many other six-column choices. result in designs that are equivalent to other six-column choices.~~

~~Even worse~~Also, the set of six-factor projections of the design in Table 5 are not exhaustive of the set of orthogonal main effects designs. It turns out that 20 of the 1,350 catalogued OAs have better A-efficiencies than the best of the six-column choices from the Plackett-Burman design. The difference can be as large as 8%, so we use the best of the catalogued designs for our comparisons rather than a sub-optimal choice of columns from the Plackett-Burman design.

Table 6: D-efficiencies and A-efficiencies for the best 24 run designs for six factors (interactions model)		
Design ID	D-efficiency (%)	A-efficiency (%)
OA #971	74	49
OA #999	74	49
OA #1031	74	49
OA #1253	79	47
OA #1259	79	47
OA #1264	79	47
OA #1332	79	47
D-optimal	92	85

In Table 6, we show the D-efficiencies and A-efficiencies of the best 24-run OAs. The best D-efficiency is 79%. There are eight arrays with this efficiency though four are not shown because they have A-efficiencies of 45%. The table also lists three designs having the maximum A-efficiency (49%).

These A-optimal OAs have a 2% better A-efficiency than four D-optimal OAs. ~~The D-optimal design obtained from JMP 11 has a D-efficiency of 92% and an A-efficiency of 85%. These efficiencies are so much higher than the OA alternatives that the difference is important in practice. To obtain comparable inference with an OA would require one to use the most efficient 28 run OA.~~

Both OAs have absolute correlation between main effects and 2FIs of either zero or one-third.

OA#1031 has 24 of 90 absolute correlations of one-third (average absolute correlation = 0.089).

OA#971 has 21 of 90 absolute correlations of one-third (average absolute correlation = 0.078). The absolute values of correlations between main effects and 2FIs are for the D-optimal design are more variable but the maximum absolute correlation is 0.1748. The average absolute correlation is 0.077.

The effect that these correlations have on the relative standard errors of the parameters is surprising. For the full interactions model the D-optimal design averages 133% more efficient than either OA over all the parameters. Individual relative efficiencies for the D-optimal design range from 116% at a minimum to a worst case of 149%.

Table 7: Designs of 24 Runs for 6 Two-level Factors																		
Run	D-optimal						OA #1332						OA #971					
1	+	+	-	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-
2	+	+	+	-	+	-	-	-	-	-	+	-	-	-	-	-	-	-
3	-	-	-	+	+	-	-	-	-	-	+	-	-	-	-	+	+	+
4	-	+	-	-	+	-	-	-	+	+	+	-	-	-	+	-	+	+
5	-	-	-	-	-	-	-	-	+	-	+	+	-	-	+	-	-	+
6	-	-	+	-	+	-	-	-	+	+	-	-	-	-	+	+	+	-
7	-	+	+	+	-	+	-	+	-	+	-	-	-	+	-	-	+	-
8	+	-	-	-	+	+	-	+	-	+	+	+	-	+	-	+	+	+
9	-	-	+	+	-	+	-	+	+	-	-	-	-	+	+	-	+	+
10	-	+	-	-	-	+	-	+	+	-	+	+	-	+	+	+	-	-
11	-	-	-	+	+	+	-	+	+	+	-	+	-	+	+	+	-	+
12	+	-	-	+	-	+	-	+	+	+	+	-	-	+	+	+	+	-
13	-	+	+	-	-	-	+	-	-	+	-	+	+	-	-	+	-	-
14	+	+	-	-	-	-	+	-	-	+	+	-	+	-	-	+	+	+
15	-	-	+	+	-	-	+	-	+	-	-	+	+	-	+	-	+	-
16	-	+	-	+	-	-	+	-	+	-	+	-	+	-	+	-	+	+
17	+	-	+	-	-	-	+	-	+	+	-	-	+	-	+	+	-	+
18	+	-	-	+	+	-	+	-	+	+	+	+	+	-	+	+	+	-
19	-	-	+	-	-	+	+	+	-	-	-	-	+	+	-	-	-	+
20	-	+	+	+	+	-	+	+	-	-	-	+	+	+	-	-	+	-
21	+	-	+	+	+	+	+	+	-	-	+	+	+	-	+	-	-	-
22	+	+	+	+	-	-	+	+	-	+	+	-	+	+	-	+	+	+
23	+	+	+	-	-	+	+	+	+	-	+	-	+	+	+	-	-	-
24	-	+	+	-	+	+	+	+	+	+	-	+	+	+	-	-	+	+

Loeppky et al. (2007) propose two sets of three six-factor designs in 24 runs, where the sets are obtained by different procedures. The designs found by either of the procedures are isomorphic to the designs found by the other procedure. We checked the designs in Table 6 against these designs. The first of their designs corresponds to OA 1332; the third one corresponds to OA 1259. We did not

include their second design, because its A efficiency is 2% smaller than the designs with D-efficiency 79% that we did include.

Table 7 shows one of the most D-efficient OAs (OA 1332; middle panel) and one of the most A-efficient OAs (OA 971; right panel). The left-most design in Table 7 is the D-optimal design from JMP [11](#). This design is not orthogonal. Like the 12-run and 20-run D-optimal designs, the columns of the 24-run D-optimal design are not balanced. This is evident by noting that there are only 10 pluses in the first column. Despite this lack of perfect balance, [the D-optimal design is substantially more efficient for the estimation of the interactions model.](#) ~~this design has a D-efficiency of 92% and an A-efficiency of 85%. These efficiencies are so much higher than the OA alternatives that the difference is important in practice. To obtain comparable inference with an OA would require one to use the most efficient 28 run OA.~~

As discussed before, OAs have D-efficiencies and A-efficiencies of 100% for models consisting of only main effects. It is natural to ask how well the D-optimal design performs for the main effects model, [since it sacrifices orthogonality of the main effects to obtain more precise estimates of 2FIs.](#) Though it is not orthogonal, the D-optimal design has a D-efficiency of 98% and an A-efficiency of 96% [for the main effects model.](#) ~~so confidence intervals for the main effects will be 2% longer on average.~~ This [difference-small loss of efficiency](#) would rarely be of practical importance.

3.4 Construction and comparison of 28 run designs for six factors

As for the 24 run designs, a commonly employed method for constructing a six-factor orthogonal array in 28 runs is to choose six factors from the Plackett-Burman design (Plackett and Burman, 1946) with 27 factors. Here, there are 296,010 ways to choose six columns out of the 27, while there

are 17,826 catalogued OAs in six factors and 28 runs. So, again, most of the six-column choices from the Plackett-Burman design are not distinct.

The best OA has a D-efficiency for the interactions model of 89% and an A-efficiency of 77%. This design outperforms the best 28-run design found by Loeppky et al. (2007). The best D-efficiency of the designs proposed by these authors is 82% and the best A efficiency is 61 %. The best design obtained with the D-optimal design algorithm has a D-efficiency ~~for the interactions model~~ of 93% and an A-efficiency of 85%. The D-efficiency and the A-efficiency for a model with only the six main effects are 99 and 98%, respectively. The best OA and the D-optimal design are shown in Table 8.

The difference in efficiencies between the D-optimal design and the best OA is less marked for the 28 run case than for the 24 run case. However, the detailed study in the next section again favors the D-optimal design, unless only three or fewer of the six factors have non-negligible effects (active factors).

Run	D-optimal						OA #17825					
1	+	+	-	-	-	-	-	-	-	-	-	-
2	+	-	+	+	+	+	-	-	-	-	-	+
3	+	-	-	+	-	-	-	-	-	-	+	-
4	-	+	-	-	-	+	-	-	-	+	-	+
5	-	+	+	+	+	+	-	-	+	-	+	-
6	+	-	-	-	+	-	-	-	+	+	-	-
7	-	+	+	-	+	-	-	-	+	+	+	+
8	-	+	-	+	+	+	-	+	-	-	+	+
9	+	+	-	+	-	+	-	+	-	+	-	+
10	+	+	+	-	-	-	-	+	-	+	+	-
11	-	+	-	+	-	-	-	+	+	-	-	-
12	+	+	+	-	+	+	-	+	+	-	-	+
13	+	+	-	+	+	-	-	+	+	+	+	-
14	+	+	-	-	+	+	-	+	+	+	+	+
15	-	+	-	-	+	-	+	-	-	-	+	+
16	+	-	+	+	-	-	+	-	-	+	-	-
17	-	-	-	-	-	-	+	-	-	+	+	-
18	+	-	-	+	+	+	+	-	+	-	-	+
19	-	-	-	+	+	-	+	-	+	-	+	-
20	+	-	+	-	+	-	+	-	+	+	-	+
21	+	+	+	+	-	+	+	-	+	+	+	+
22	-	-	-	-	+	+	+	+	-	-	-	+
23	-	-	+	-	-	+	+	+	-	-	+	-
24	+	-	-	-	-	+	+	+	-	+	-	-
25	-	+	+	+	-	-	+	+	-	+	+	+
26	-	-	-	+	-	+	+	+	+	-	-	-
27	-	-	+	+	+	-	+	+	+	-	+	+
28	+	+	+	+	+	-	+	+	+	+	-	-

4 Model Spaces with Subsets of Factors or Interactions

In this section, we compare the performance of our six-factor designs for model spaces previously introduced in the design literature. The first, which we call the projection model space, was considered by Loeppky et al. (2007). Models in the projection model space consist of subsets of factors and all two-factor interactions involving these factors. For example, there are 20 three-factor models in the projection model space for the six factor designs. These 20 models consider all combinations of six factors taken three at a time. Each model consists of the main effects of the three specified factors and their three two-factor interactions. We compare designs for their average estimation performance across all two-factor, three-factor, four-factor and five-factor projections.

We also consider the main effects plus interactions model space (MEPI) also introduced by Li and Nachtsheim (2000). MEPI models have all the main effects and up to g additional interactions, where $g = 1, 2, \dots, n(n-1)/2$. In our comparisons we only consider the addition of two to five additional 2FIs. To compare the MEPI model space for four additional 2FIs, our models start with all six main effects. Then, we add all possible combinations four of the six possible combinations (10) of 2FIs.

We make these comparisons for the six factor case only because we define screening designs as having six or more factors. A major assumption in screening is that investigators do not expect all the factors to have strong effects. If so, it makes sense to compare designs based on models that involve only a subset of the factors.

4.1 Six Factor Designs with 24 runs

The plots in Figure 1 show the averages (heavy dots) and ranges (“error bars”) of D-efficiencies for interactions models in 2-5 factors based on the three 24-run designs from Table 7. The designs are the D-optimal design, OA 1332, and OA 971, respectively. We prefer OA 1332 over the other OAs having the same D-efficiency for the interactions model in six factors, because it has slightly better

D-efficiencies for both the projection and MEPI model spaces. OAs 971 and 999 both have the same A-efficiencies in these models; these are slightly better than OA 1031. Our decision to plot OA 971 instead of OA 999 was arbitrary.

The plots in the four groupings of the figure are based on 15, 20, 15 and 6 different models, because these are the numbers of different projections into 2, 3, 4 and 5 factors, respectively. Except for the two-factor projections, the D-efficiency for the A-optimal OA, 971, is worse than one or both of the other two designs.

Figure 1. D-Efficiency of 2 – 5 factor Projections for Three Designs with Six Factors and 24-Runs
Designs

~~The plots in the four groupings of the figure are based on 15, 20, 15 and 6 different models, because these are the numbers of different projections into 2, 3, 4 and 5 factors, respectively. Except for the two factor projections, the D-efficiency for the A-optimal OA, 971, is worse than one or both of the other two designs.~~

If it were possible to know that the maximum number of active factors would not exceed three, we would prefer the most D-efficient orthogonal array (OA 1332) to the D-optimal design. For OA 1332 16 of the 20 projections into three factors have a better D-efficiency than the best three-factor projections for the D-optimal design. On the other hand, four projections have slightly worse efficiencies than the worst projection of the D-optimal design.

For projections into four and five factors, the superiority of the D-optimal design is clear. For four active factors, the worst projection of the OA 1332 is 5% less efficient than the worst projection for the D-optimal design, while the remaining projections of the orthogonal array are only 2% better than the worst projection of the optimal design. We are generally uncomfortable risking the success of a study on the possibility of an unlucky break in which a set of four or five factors turns out to be active.

The plots in Figure 2 mimic those of Figure 1 except that they show A-efficiencies rather than D-efficiencies. show the averages (heavy dots) and ranges (error bars) of the A-efficiencies of the three selected designs. It is remarkable that for the 5 factor projections, the worst A-efficiency of OA 1332 is as good as the best A-efficiency of the most A-efficient orthogonal array (OA 971). The general picture is the same as for the D-efficiencies: for interactions models in up to three factors, the D-optimal design is slightly worse than the best orthogonal array. As in the study of D-efficiencies, the D-optimal design performs better in A-efficiency for projections to four or five factors. If it were possible to know that only three or fewer factors were going to be active, we would prefer OA 1332.

Figure 2. A-Efficiencies for Projections of 2 – 5 Factors for Three Designs with Six Factors and 24 Runs

Figure 3 shows the three designs' efficiencies with respect to MEPI models with 2–5 of the 15 two-factor interactions. There are $105 = (15 \times 14) / 2$ models in the MEPI model space for 2 factors. For the 3–5 factor model spaces we consider 455, 1365 and 3003 models respectively. The most striking feature of Figure 3 is the substantial variation in efficiencies for the OAs as opposed to those for the optimal design. For example, the A-efficiencies for models with five two-factor interactions only range from 91% to 93% for the D-optimal design. However, for OA 1332 the A-efficiencies have a far wider range (86% to 98%). The A-efficiencies of OA 971 have an even wider range (70% to 96%). Since an investigator cannot know in advance which or how many two-factor interactions will be active, for the MEPI model space the D-optimal design appears to be the least risky choice.

Figure 3. A-Efficiencies for Different 24 Run Designs and MEPI Models with 2 – 5 Interactions

In summary, an attractive feature of the D-optimal design is the fact that the efficiencies shown in Figure 1– 3 are both high and relatively homogeneous. This implies that the design behaves well no matter which subset of factors or interactions is active. This feature is not shared by the best OAs.

4.2 Designs with 28 runs

For 28 runs, we compare the optimal design with OA 17825, which is the most D-efficient as well as the most A-efficient orthogonal array with six factors; the designs were given in Table 8. Figure 4 shows the A-efficiencies for the projection model spaces involving 2 to 5 factors for both the D-

optimal design and OA 17825. Note that all the sub-models for the OA have the same A-efficiency within each projection. The A-efficiencies for two-factor projections is 100%, which is a characteristic of strength 2 OAs. As the dimensionality of the projection increases, the A-efficiency of the OA drops from 100% for 2 factor projections to 87% for 5 factor projections.

The A-efficiencies for 2 factor projections of the D-optimal design range from 96% to 99%. For 4 and 5 factor projections, all the projections of the D-optimal design have higher A-efficiencies than the A-efficiencies of OA #17825. As for the 24 run case, we would prefer the OA if it were possible to know that there are no more than three active factors. Otherwise, given uncertainty about the number of active factors, we prefer the estimation performance of the D-optimal design in projection model spaces with more than three factors.

Figure 4. A-Efficiencies for Projections of 2 – 5 Factors for Two 28 Run Designs

Figure 5 shows the A-efficiencies for MEPI models having 2 to 5 added 2FIs. There are 105 models containing only two 2FIs in addition to the main effects. For all of these models OA #17825 has the same A-efficiency (97.8%). For the same case, the D-optimal design A-efficiencies ranging from 96.4% to 97.2%. As the number of added 2FIs grows from 2 to 5, the A-efficiency of the D-optimal design exceeds the A-efficiency of the OA to an increasing degree.

Figure 5. A-Efficiencies for Different 28 Run Designs and MEPI Models with 2 – 5 Interactions

5 Discussion

This paper features the comparison of two-level strength-2 OAs with computer-generated optimal design alternatives. For cases with several two-level factors and one or more multi-level factors, we refer to Schoen (2010), who compared strength-3 OAs with optimal designs to estimate the interactions model.

We investigated alternatives to the standard candidate designs for four to six factors for fitting the interactions model. We also studied the performance of the six-factor designs for the projection and MEPI model spaces. Any of the alternative designs we studied could prove useful in cases where the cost of performing 32 runs is prohibitive. We focused on designs of 24 and 28 runs because both OAs and D-optimal designs are available for these cases. The optimal design approach ~~is preferable~~ yields more precise parameter estimates for interactions models having four or more factors and for MEPI models with three or more active interactions in addition to the main effects. For cases where the OAs perform better the difference between the design approaches is too small to be practically important.

In case even the cost of 24 runs is prohibitive to investigate six factors, we note the availability of the 22-run design given by Rechtschaffner (1967). This design is D-optimal for the interactions model but saturated (i.e. there are no degrees of freedom left to estimate error). The D-efficiency for the interactions model is 92.6 % and the A-efficiency is 86.8 %. These efficiencies are higher than those of the D-optimal 24-run design considered here. To investigate this issue further, we augmented the Rechtschaffner design in several ways to obtain a 24-run design. The utility of our 24 run D-optimal design was confirmed, because the augmented designs had worse efficiencies for the interactions model as well as for the various submodels studied in this paper.

The D-optimal Rechtschaffner design illustrates the fact that budgeting runs for optimal designs is flexible. It is possible to generate optimal designs for any run budget greater than the number of required parameters to be fit. So, designs of 26 or 30 runs are also feasible and can provide more degrees of freedom for estimating the error variance than the 24 or 28 run designs respectively. Of course, the fact that the 2^{6-1} fractional factorial design is orthogonal and offers minimum variance estimates for all main effects and 2FIs means~~has so many desirable features~~ that unless runs are expensive, it is hard to imagine choosing a 30 run design instead.

Our research allows us to make a number of recommendations for the practitioner. First, it is common for investigators to create designs from a tabled orthogonal array by using a small set of arbitrarily chosen columns from an array having many more columns that are required. ~~This is especially true for studies using the Taguchi arrays.~~ If one knew that only the main effects were active, then this would have no undesirable consequences. But, this is exactly what one does not know in advance of experimentation. This paper shows that creating designs by choosing columns from a Plackett-Burman design can lead to ~~a couple of~~ undesirable outcomes. ~~One may be unable to fit models that it should be possible to fit or one may be able to fit the required model but with very low efficiency (or high variance inflation).~~ First, if many 2FIs are active, it may not be possible to estimate them. Second, even if all the active 2FIs are estimable, the variances of these estimates may be unnecessarily high.

This paper also calls into question the commonly made assertion that D-optimal designs do not have good projection properties. In our examples we see that D-optimal designs are outperforming the most efficient OAs for fitting interactions models in projections of six factors down to four factors.

Practitioners need to understand that strength-2 OAs only guarantee the orthogonality of the main effects. For these OAs, the 2FIs are generally not orthogonal to each other. The degree of nonorthogonality as evidenced by higher standard errors among the 2FIs in these OAs exceeds that of the D-optimal design. The D-optimal design for the interactions model trades off orthogonality of the main effects in exchange for better estimation of 2FIs than any strength-2 OA can provide. For fitting only the main effects, the nonorthogonality of the six-factor D-optimal design for the interactions model in 24 runs leads to a small loss of precision in the estimates of these effects that is maximum VIF of 1.04. ~~Such a small variance inflation is~~ negligible in almost any practical application.

Finally, practitioners should be aware that it is not necessary to allocate the resources required by a resolution V fractional factorial design to fit the interactions model. ~~The 24-run choices we recommend yield a 25% savings in resource cost versus the standard 2^{6-1} fractional factorial design.~~

There is scope for further research along these lines. While we have investigated only cases with four to six factors, our methodological approach is general. For example, the smallest resolution V fractional factorial design for seven factors requires 64 runs, which is seldom a feasible choice. However, the interactions model only has 29 parameters, so it is likely that strength 2 OAs of 32, 36 or 40 runs can fit these models. So far, catalogs of OAs for designs of this size are unavailable due to the magnitude of the computing problem and the huge numbers of OAs.

References

Atkinson, A. C., Donev, A. N. and Tobias, R. D. (2007). *Optimum experimental designs, with SAS*. (Oxford: Oxford University Press).

~~Candes, E. and Tao, T. (2007). "The Dantzig Selector: Statistical Estimation When p Is Much Larger Than n ". *Annals of Statistics* 35, pp. 2313–2351.~~

~~Daniel, C. (1959) Use of half-normal plot in interpreting factorial two-level experiments," *Technometrics*, 8:177.~~

~~Goos, P. and Jones, B., (2011) *Optimal Design of Experiments A Case Study Approach*, Wiley Ltd. Chichester, U.K.~~

Hedayat, A. S., Sloane, N. J. A. and Stufken, J. (1999) *Orthogonal Arrays: Theory and Applications* (New York: Springer).

~~Hurvich, C. M. and Tsai, C. L. (1989). "Regression and Time Series Model Selection in Small Samples". *Biometrika* 76, pp. 297–330.~~

John, P.W.M. (1962) "Three quarter replicates of 2^n designs". *Biometrics*, 18: 172-184.

JMP Software version 11 (2013) SAS Institute, Cary, North Carolina

Li, W., and Nachtsheim, C. J. (2000) "Model-Robust Factorial Designs". *Technometrics*, 42: 345-352.

Loeppky, J., Sitter, R.R. and Tang, B. (2007) "Nonregular designs with desirable projection properties". *Technometrics*, 49: 454-467.

Mee, R.W. (2004) "Efficient two-level designs for estimating all main effects and two-factor interactions". *Journal of Quality Technology*, 36:400-412.

Meyer, R. K., and Nachtsheim, C. J. (1995) "The Coordinate Exchange Algorithm for Constructing Exact Optimal Experimental Designs". *Technometrics*, 37:60-69.

Plackett, R. L. and Burman, J. P. (1946) "The Design of Optimum Multifactorial Experiments". *Biometrika*, 33:305-325.

Rechtschaffner, R.L. (1967) Saturated fractions of 2^n and 3^n factorial designs. *Technometrics*, 9:569-575.

Schoen, E. D., Eendebak, P. T., and Nguyen, M. V. M. (2010) "Complete Enumeration of Pure-Level and Mixed-Level Orthogonal Arrays". *Journal of Combinatorial Designs*, 18:123-140.

Schoen, E.D. (2010) "Optimum designs versus orthogonal arrays for main effects and two-factor interactions". *Journal of Quality Technology*, 42: 197-208.

~~Tibshirani, R. (1996) "Regression shrinkage and selection via the lasso". *J. R. Statist. Soc. B*, **58**, 267-288.~~

~~Zou, H. and Hastie, T. (2005) "Regularization and variable selection via the elastic net", *Journal of the Royal Statistical Society Series B* 67 Part 2 pp. 301-320.~~