DEPARTMENT OF ENGINEERING MANAGEMENT

Orthogonal blocking arrangements for 24-run and 28-run two-level designs

Eric D. Schoen, Nha Vo-Thanh & Peter Goos

UNIVERSITY OF ANTWERP
Faculty of Applied Economics
City Campus
Prinsstraat 13, B.226
B-2000 Antwerp
Tel. +32 (0)3 265 40 32
Fax +32 (0)3 265 47 99
www.uantwerpen.be

FACULTY OF APPLIED ECONOMICS

DEPARTMENT OF ENGINEERING MANAGEMENT

Orthogonal blocking arrangements for 24-run and 28-run two-level designs

Eric D. Schoen, Nha Vo-Thanh & Peter Goos

RESEARCH PAPER 2016-002
FEBRUARY 2016

University of Antwerp, City Campus, Prinsstraat 13, B-2000 Antwerp, Belgium
Research Administration – room B.226
phone: (32) 3 265 40 32
fax: (32) 3 265 47 99
e-mail: joeri.nys@uantwerpen.be

D/2016/1169/002

Orthogonal blocking arrangements for 24-run and 28-run two-level designs

Eric D. Schoen
Faculty of Applied Economics, University of Antwerp, Belgium
TNO, Zeist, Netherlands

Nha Vo-Thanh
Faculty of Applied Economics, University of Antwerp, Belgium

Peter Goos
Faculty of Applied Economics and StatUa Center for Statistics, University of Antwerp, Belgium
Faculty of Bioscience Engineering and Leuven Statistics Research Centre (LSTAT), KU Leuven, Belgium

February 11, 2016

Abstract

Much research has been done recently concerning 24-run two-level designs involving 3–23 factors and 28-run two-level designs involving 3–27 factors. The focus of this research was on completely randomized screening designs, and led to lists of recommended designs for each number of factors. For the recommended designs, the aliasing between main effects and two-factor interaction effects is either nonexisting or very limited, and, subject to this, the aliasing among two-factor interaction effects is minimized. It is, however, unclear which 24-run and 28-run designs are best when complete randomization is infeasible and the designs have to be arranged in blocks. In this paper, we address this issue and present the best arrangements of 24-run designs in 3, 4 or 6 blocks, and the best arrangements of 28-run designs in 7 blocks.

KEY WORDS: Confounding Frequency Vector; G-Aberration; G_2-Aberration; Generalized Word-Length Pattern; Orthogonal Array; Orthogonal Blocking.

1 Introduction

Blocking is one of the most important principles of experimental design. To illustrate the relevance of blocking in modern factorial screening experiments, consider the following example. In cheese production, raw milk is processed into curds in large production tanks. Many controllable factors, such as the rotor speed, the amount of rennet and the milk temperature, impact the curds production. Dairy technologists are interested in the effects of nine of these controllable curds production factors on the properties of the cheeses made from the curds. The technologists are primarily interested in the factors' main effects, but they think that a few two-factor interactions might also be active. The traditional type of experimental design to use in that scenario is a two-level orthogonal screening design.

Two-level orthogonal designs are popular experimental plans for screening, because they yield main effect estimates with maximum precision in case a model is estimated with main effects only. The maximum precision is due to the facts that the two-level orthogonal designs are balanced (each level of every factor is used equally often) and that the main effect contrast vectors are orthogonal to each other.

A complication when performing the cheese production experiment is that one batch of raw milk allows for at most eight curds to be produced and that substantial batch-to-batch variation may be present. As a result, the required design is a two-level screening design arranged in blocks of eight runs, where the blocks correspond to the batches of raw milk. Since the budget for the study allows for 24 runs, three blocks of eight experimental runs are required. Note that the original nine factors are generally called the treatment factors, while the batch is generally called the blocking factor of the experiment. Therefore, in the cheese production study, there are nine two-level treatment factors and a three-level blocking factor.

Ideally, the estimates of the treatment factors' main effects are not affected by the batch-to-batch or block-to-block variation and the estimation of the two-factor interaction effects is affected as little as possible. We call blocking arrangements for which the main effect estimates are not affected by the block-to-block variation orthogonally blocked designs. For such designs, the main effect contrast vectors are orthogonal to the blocks.

Now, it turns out that 71,157,023 different orthogonally blocked 24-run designs exist with nine two-level treatment factors and one three-level blocking factor (Schoen et al., 2010). Based on the research described below, we identified three attractive orthogonally blocked designs, all of which are shown in Table 1. The orthogonally blocked nature of the designs is due to the fact that the main effect contrast vectors of each of the nine treatment factors sum to zero in each block for each design in Table 1.

For designs 1 and 3 in Table 1, the main effects are not only orthogonal to the blocks, but also to the two-factor interactions. Therefore, these designs allow the main effects to be estimated independently from the block effects as well as from the two-factor interactions. Design 2 in Table 1 does not possess this attractive property, but it possesses another interesting characteristic: it allows 14 two-factor interaction effects to be estimated, while designs 1 and 3 only allow 11 two-factor interaction effects to be quantified. Therefore, active interactions might be detected more easily with design 2 than with designs 1 and 3.

1.1 State of the art

There is a vast amount of literature on blocking two-level designs. Most of it deals with the regular blocking of regular two-level designs. Any n-factor regular two-level design consists of a 2^p full factorial design in p basic factors. The settings of the $n - p$ remaining factors are calculated by taking products of the settings of two or more basic factors. Due to this construction, any pair of factorial effects (for instance, main effects and two-factor interactions) is either completely aliased or orthogonal (Montgomery, 2009; Wu and Hamada, 2009). Also, by construction, the run size of a regular two-level design is a power of two. Consequently, regular orthogonal blocking arrangements of regular two-level designs also have a number of runs that is a power of two, and they result in factorial effects either being completely aliased with the blocks or being orthogonal to them. References discussing orthogonally blocked regular two-level designs include National Bureau of Standards (1957), Bisgaard (1994), Sun et al. (1997), Sitter et al. (1997), Cheng and Wu (2002) and Xu (2006).

All possible orthogonal two-level designs with up to 24 runs are known and have been studied in detail. References for the designs with up to 20 runs include Deng and Tang (2002), Li et al. (2004), Xu and Deng (2005) and Sun et al. (2008). For designs involving 24 or 28 runs, results based on incomplete catalogs were given by Deng and Tang (2002), Ingram and Tang (2005), Loeppky et al. (2007), Angelopoulos et al. (2007) and Evangelaras et al. (2007). However, Schoen et al. (2010) enumerated all possible 24-run two-level designs. Schoen et al. (2015) and Bulutoglu and Ryan (2015) explored all these designs and identified the best few designs that minimize the amount of aliasing between the main effects and the two-factor interaction effects and, subject to this, the amount of aliasing among the two-factor interaction effects. They also enumerated the best 28-run designs involving 3–14 factors according to these criteria. The papers differ in focus: the first one is on design exploration, while the second one is focussed on design enumeration algorithms.

Cheng et al. (2004) were the first to propose criteria as well as a search procedure to identify good orthogonal blocking arrangements of two-level nonregular designs in case the number of blocks is a power of two. However, the blocking arrangements shown in Table 1 for the cheese production experiment involve three blocks, and can therefore not be found with the approach of Cheng et al. (2004). Schoen et al. (2013) proposed a more general search procedure for arranging factorial designs of resolution III in any number of blocks that allows orthogonal blocking. The method, which works for factors with any numbers of levels, is based on complete catalogs of orthogonal designs that include the treatment factors and the blocking factor. Schoen et al. (2013) presented the best blocking arrangements for mixed-level and pure-level designs with 12, 16, 18, 20, and 27 runs. They did not address 24-run and 28-run designs, because the complete catalogs for these designs are very large and a complete exploration of the catalogs was computationally infeasible at the time.

Table 1: Three 24-run designs with nine two-level treatment factors arranged in three blocks of size eight. The three blocks are separated by a blank line.

Design 1									Design 2									Design 3								
−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−
−	−	−	−	+	+	+	+	+	−	−	−	−	−	+	+	+	+	−	−	−	−	−	−	+	+	+
−	−	+	+	−	−	−	+	+	−	−	−	+	+	−	−	−	+	−	−	−	+	+	+	−	+	+
−	+	−	+	−	+	+	−	−	−	+	+	−	+	−	+	+	−	−	−	+	−	+	+	−	−	−
+	−	+	−	+	−	+	−	−	+	−	+	+	−	+	−	+	−	+	+	−	+	−	−	+	+	+
+	+	−	+	+	−	−	−	+	+	+	−	+	+	+	+	+	+	+	+	+	−	−	−	+	−	−
+	+	+	−	−	+	−	+	−	+	+	+	−	+	+	−	−	−	+	+	+	+	+	+	−	−	−
+	+	+	+	+	+	+	+	+	+	+	+	+	−	−	+	−	+	+	+	+	+	+	+	+	+	+
−	−	−	+	+	−	+	−	+	−	−	−	−	+	+	−	+	−	−	−	+	+	−	+	+	−	+
−	−	+	−	−	+	+	+	−	−	−	+	+	−	−	+	+	+	−	+	−	−	+	+	+	+	−
−	+	−	−	+	−	−	+	−	−	+	−	+	−	+	+	−	−	−	+	−	+	−	+	+	−	−
−	+	+	−	+	+	−	−	+	−	+	+	+	+	+	−	+	+	−	+	−	+	+	−	−	−	+
+	−	−	+	+	+	−	+	−	+	−	+	−	−	−	+	−	−	+	−	+	−	−	+	+	+	−
+	−	+	+	−	+	+	−	+	+	−	+	−	+	+	+	−	+	+	−	+	−	+	−	−	+	+
+	+	−	−	−	−	+	+	+	+	+	−	−	−	−	−	+	+	+	−	+	+	−	−	−	−	+
+	+	+	+	−	−	−	−	−	+	+	−	+	+	−	−	−	−	+	+	−	−	+	−	−	+	−
−	−	+	+	+	+	−	−	−	−	−	+	+	+	+	+	−	−	−	−	+	+	+	−	+	+	−
−	+	−	+	−	+	−	+	+	−	+	−	−	+	−	+	−	+	−	+	+	−	−	+	−	+	+
−	+	+	−	−	−	+	−	+	−	+	+	−	−	+	−	−	+	−	+	+	−	+	−	+	−	+
−	+	+	+	+	−	+	+	−	−	+	+	+	−	−	−	+	−	−	+	+	+	−	−	−	+	−
+	−	−	−	−	+	−	−	+	+	−	−	+	−	+	−	−	+	+	−	−	−	+	+	+	−	+
+	−	−	+	−	−	+	+	−	+	−	−	+	+	−	+	+	−	+	−	−	+	−	+	−	+	−
+	−	+	−	+	−	−	+	+	+	−	+	−	+	−	−	+	+	+	−	−	+	+	−	+	−	−
+	+	−	−	+	+	+	−	−	+	+	−	−	−	+	+	+	−	+	+	−	−	−	+	−	−	+

1.2 Purpose and organization

The purpose of the present paper is to find optimal orthogonal blocking arrangements for 24-run and 28-run orthogonal designs. These designs are nonregular, because the run sizes 24 and 28 are not powers of two. Most pairs of factorial effects from such designs are not completely aliased. For this reason, nonregular designs permit larger numbers of statistical models to be fitted to the data, thereby offering more information on the factorial effects. In addition, in case certain interactions are active and a main effects model is estimated, the estimates of the main effects usually have a smaller bias.

The ever increasing computing power permitted us to find optimal blocking arrangements of 24-run designs in three, four, or six blocks and optimal blocking arrangements of 28-run designs in seven blocks by exploring complete catalogs of nonregular two-level designs with an additional 3-, 4-, 6- or 7-level factor. To deal with specific features of the 24-run and 28-run designs, we slightly adapted the general optimality criteria of Schoen et al. (2013).

The remainder of the paper is organized as follows. In Section 2, we review optimality criteria for completely randomized and blocked orthogonal two-level designs. Optimal blocking arrangements of the 24-run and 28-run designs are presented in Section 3. Finally, in Section 4, we discuss our findings and provide a few suggestions for future research.

2 Optimality criteria

In this section, we first review optimality criteria for completely randomized two-level designs. Next, we explain how to quantify the confounding and aliasing in blocked designs. We propose three criteria to deal with blocked orthogonal two-level designs and we illustrate these criteria using the three designs for the cheese making experiment in Table 1.

2.1 Criteria for completely randomized designs

Quality criteria for two-level orthogonal designs in the absence of blocking include the strength of the design (Rao, 1947), the rank of the interaction model matrix (which corresponds to a model with an intercept, all main effects and all two-factor interactions; Cheng et al., 2008), G-aberration (Deng and Tang, 1999, 2002), and G_2-aberration (Tang and Deng, 1999).

2.1.1 Strength

Orthogonal two-level designs have a strength of t if, for any set of t factors, all 2^t combinations of factor levels occur equally often. When using strength-2 designs, the main effects are orthogonal. When using strength-3 designs, the main effects are not only orthogonal to each other, but also to the two-factor interaction effects. When using strength-4 designs, the two-factor interactions are orthogonal to each other too. A consequence of the definition of strength is that any strength-t design has a run size proportional to 2^t. The 24-run designs studied in this paper can therefore either have strength 2 or strength 3. The 28 run designs all have a strength of 2.

2.1.2 Rank

The rank of the interaction model matrix is of interest to researchers who want to identify active two-factor interactions in addition to main effects. The larger the rank, the larger the number of estimable two-factor interactions. For example, the rank of the interaction model matrix of the first design option in Table 1 equals 21. The second option has a rank of 24. The larger rank might be a reason to prefer the second option over the first.

2.1.3 G-aberration

The G-aberration criterion is based on counting correlations between contrast vectors. Of particular interest in orthogonal two-level designs are correlations between main effect contrast vectors and two-factor interaction contrast vectors (type-3 correlations) and correlations among the two-factor interaction contrast vectors (type-4 correlations). Deng and Tang (1999) showed that the only possible absolute type-3 and type-4 correlations for orthogonal N-run two-level designs equal $(N - 8j)/N$, where j is a non-negative integer smaller than or equal to $N/8$. So, the only possible absolute type-3 and type-4 correlations are 1, 2/3, 1/3 and 0 for orthogonal 24-run designs, and 1, 5/7, 3/7 and 1/7 for orthogonal 28-run designs. For any given design, the F_3 vector quantifies with which frequencies the type-3 correlations occur for pairs of one main effect and one two-factor interaction effect. The F_4 vector quantifies with which frequencies the type-4 correlations occur for pairs of two two-factor interaction effects. More specifically, the F_3 and F_4 vectors contain the frequencies with which the different type-3 and type-4 correlations occur divided by 3. The division by 3 for the type-3 correlations is due to the fact that, for any set of three factors, there are three pairs of one main effect contrast vector and one two-factor interaction contrast vector, all of which have the same correlation. The frequencies for the type-4 correlations are divided by 3 for a similar reason. Generally, the first entries for the F_3 and F_4 vectors correspond to the most severe correlations, while the last entries correspond to the least severe correlations. The frequencies of the zero type-3 and type-4 correlations are usually omitted from the F_3 and F_4 vectors. Finally, in the notation of the F_3 and F_4 vectors, it is customary to refer to the inner products between the contrast factors rather than to their correlations. So, an F_3 vector for a 24-run design is denoted by $F_3(24, 16, 8)$, while an F_3 vector for a 28-run design is denoted by $F_3(28, 20, 12, 4)$. A similar notation is used for the F_4 vectors.

For an n-factor design, in addition to the F_3 and F_4 vectors, one can also define the $F_5, \ldots, F_n$ vectors involving the frequencies of the different absolute type-5, ..., type-n correlations. Together, all these vectors form the confounding frequency vector $F_3, F_4, \ldots, F_n$, which is usually abbreviated as CFV. The G-aberration of a design is its ranking according to the CFV; a minimum G-aberration design minimizes the CFV's entries from left to right.

Note that the F_5 and F_6 vectors quantify the extent to which three-factor interaction contrast vectors are correlated with two-factor interaction contrast vectors and other three-factor interaction contrast vectors, respectively. In most applications, three-factor interactions are assumed to be negligible. Therefore, it is not particularly interesting to minimize the F_5 and F_6 vectors. The same is true for the $F_7, \ldots, F_n$ vectors, which deal with interactions of an even higher order. Therefore, in this paper, we restrict our attention to the F_3 and F_4 vectors. Note that strength-3 designs do not involve type-3 correlations. Therefore, the elements of their F_3 vectors all equal 0.

2.1.4 G_2-aberration

Tang and Deng (1999) proposed a simpler version of the CFV called the generalized word length pattern (GWLP). To calculate the GWLP's first element, all type-3 correlations are squared and summed, and the result is divided by 3. The resulting figure is called the generalized word count of length 3 and denoted by A_3. Generalized word counts of length $4, 5, \ldots, n$ are defined in a similar fashion, and denoted by $A_4, A_5, \ldots, A_n$. The GWLP is the vector $(A_3, A_4, \ldots, A_n)$. The G_2-aberration of a design is its ranking according to the GWLP; a minimum G_2-aberration design minimizes the GWLP's entries from left to right. We restrict our attention to the A_3 and A_4 counts for reasons similar to those given in Section 2.1.3. Finally, note that, for strength-3 designs, $A_3 = 0$.

2.1.5 Illustration

For the first and third 24-run design options in Table 1, the F_3 and F_4 vectors equal $F_3(24, 16, 8) = (0, 0, 0)$ and $F_4(24, 16, 8) = (0, 0, 126)$, respectively. The second design option has $F_3(24, 16, 8) = (0, 0, 10)$ and $F_4(24, 16, 8) = (0, 0, 98)$. So, when using the first option, there are no occurrences of the absolute type-3 correlations of 1, 2/3 and 1/3 between a main effect contrast vector and a two-factor interaction contrast vector. In other words, all main effect contrast vectors are orthogonal to the two-factor interaction contrast vectors. Moreover, the only occurring absolute type-4 correlation is 1/3, meaning that every pair of two-factor interaction effect estimates has a correlation of $+1/3$ or $-1/3$.

When the second 24-run design option would be used, $10 \times 3 = 30$ pairs of a main effect contrast vector and a two-factor interaction contrast vector would have a correlation of $\pm 1/3$. This can be seen from the value of 10 in the F_3 vector for that design. The remaining pairs of a main effect contrast vector and a two-factor interaction contrast vector would be uncorrelated.

Starting from the F_3 and F_4 vectors, the A_3 and A_4 values of the design options in Table 1 can be computed. For the first and third design options, $A_3 = 0 \times 1^2 + 0 \times (2/3)^2 + 0 \times (1/3)^2 = 0$, and $A_4 = 0 \times 1^2 + 0 \times (2/3)^2 + 126 \times (1/3)^2 = 14$. For the second design option, $A_3 = 1.11$ and $A_4 = 10.89$.

2.2 Confounding and aliasing in blocked designs

When screening designs have to be blocked, a primary concern is to ensure that the main effects can be estimated independently from the block effects. For this reason, we restrict our attention to orthogonally blocked designs. The next most important concern involves the confounding of the two-factor interactions with the blocks. By far the most common approach is to minimize the confounding of the two-factor interactions with the blocks. In doing so, it might be easier to detect active interaction effects. Schoen et al. (2013), however, explain that maximizing the confounding between the two-factor interactions and the blocks may be sensible as well. As a matter of fact, this alternative approach would ensure that the mean squared error of the main effects model has the smallest possible upward bias due to active interactions, and would result in a maximum power for detecting active main effects when using a model involving main effects only.

Each of the two approaches to deal with two-factor interactions requires measuring the extent to which they are confounded with the block effects. Schoen et al. (2013) quantified the confounding of the two-factor interactions with the blocks by means of the FA_3 vector, a frequency vector of generalized counts of length-3 words that involve the blocking factor.

To explain the construction of the FA_3 vector, consider the problem of arranging a two-factor two-level treatment design with 24 runs into three blocks. The first step is to define an orthogonal coding for the three-level blocking factor. One possible choice of coding involves dummy variables c_1 and c_2 which takes the values -1 and 1, respectively, for the first block, 0 and -2 for the second block, and 1 and 1 for the third block. The next step is to normalize the contrast vector containing all 24 c_1 and c_2 values to length $\sqrt{24}$ and to calculate the inner products between the treatment factors' two-factor interaction contrast vector, on the one hand, and the normalized contrast vectors with c_1 and c_2 values, on the other hand. Next, the two resulting inner products have to be divided by 24. This results in two correlations between the interaction contrast vector and the contrast vectors of the three-level blocking factor. Finally, summing the squared correlations produces the generalized word count of length 3 for the $3 \times 2 \times 2$ design involving the three-level blocking factor and the two two-level treatment factors. We denote that generalized word count of length 3 by $A_3^{(b22)}$. Xu and Wu (2001) show that the $A_3^{(b22)}$ value does not depend on the exact choice of the orthogonal coding scheme for the blocking factor.

The smaller the value of $A_3^{(b22)}$, the less severe the aliasing between the main effect of the three-level blocking factor and the interaction involving the two-level treatment factors. There are four different orthogonal 24-run designs with one three-level factor and two two-level factors (Schoen et al., 2010). Studying these four designs shows that the possible generalized word counts of length 3 for one three-level factor and two two-level factors are 2/3, 1/2, 1/6 and 0, respectively. In other words, the only possible values of $A_3^{(b22)}$ are 2/3, 1/2, 1/6 and 0. Smaller values indicate a smaller degree of confounding between the two-factor interactions and the blocks. A zero value for this count indicates orthogonality between the two-factor interaction and the blocks.

For any orthogonal 24-run design with one three-level (blocking) factor and n two-level (treatment) factors, there exist $n(n-1)/2$ projections of the design into one three-level factor and two two-level factors. The FA_3 vector lists the frequencies with which the nonzero $A_3^{(b22)}$ values of 2/3, 1/2 and 1/6 occur in these projections. For example, the first two designs in Table 1 have an FA_3 vector of (0, 0, 18). This means that 18 of the 36 two-factor interactions have an $A_3^{(b22)}$ value of 1/6 (indicating some confounding of the two-factor interactions with the blocks), while the remaining 18 have an $A_3^{(b22)}$ value of zero (indicating no confounding with the blocks). For the third design option in Table 1, the FA_3 vector is (0, 9, 18). As a result, when that design would be used, nine additional interactions would be confounded with the blocks. The extent of that additional confounding would also be more severe, because of the $A_3^{(b22)}$ value of 1/2 for these nine interactions.

A more concise measure of the confounding between the two-factor interactions and the blocks is the total generalized count of length-3 words involving the blocking factor and the two-factor interaction contrast vectors, which we call the (total) block count and denote by $A_3^{(btot)}$. We obtain this count by multiplying the frequencies in the FA_3 vector with the corresponding generalized word counts and summing the resulting products. For example, the block count $A_3^{(btot)}$ of the first two designs in Table 1 equals $0 \times 2/3 + 0 \times 1/2 + 18 \times 1/6 = 3$, whereas it equals 7.5 for the third design. Obviously, the smaller the block count, the smaller the overall extent to which two-factor interactions are confounded with the blocks.

2.3 Criteria for optimal blocking

To identify optimal blocking arrangements, we adapt the criteria proposed by Cheng and Wu (2002) for regular two-level designs, Cheng et al. (2004) for nonregular orthogonal two-level designs and Schoen et al. (2013) for general orthogonal nonregular designs to the specific cases of blocked nonregular two-level designs considered here. We name the resulting criteria the W_2, W_2^- and W_3 criteria and we collectively call them 'the three criteria'. Like the G-aberration and G_2-aberration criteria, these criteria involve a vector that has to be minimized from left to right. The difference between the aberration criteria and our criteria is that our minimization vectors involve mixes of the generalized word counts A_3, A_4 and $A_3^{(btot)}$, the F_3, F_4 and FA_3 vectors, and, in one criterion, the rank R of the interaction model matrix. The exact minimization vectors for our design selection criteria are shown in Table 2.

Table 2: Three criteria for selecting optimally blocked two-level designs

Criterion	Minimization vector					
W_2	$-R$	A_3	F_3	$A_3^{(btot)}$	A_4	F_4
W_2^-	A_3	F_3	$-A_3^{(btot)}$	A_4	F_4	
W_3	F_3	FA_3				

2.3.1 The W_2 criterion

As can be seen from the first element of its minimization vector, $-R$, our W_2 criterion seeks designs with a large rank R for the interaction model matrix, i.e., designs that allow many two-factor interaction effects to be estimated. The minus sign in $-R$ indicates that minus the rank should be minimized, which is the same as maximizing the rank itself. The inclusion of $-R$ as the first element in the W_2 minimization vector distinguishes our W_2 criterion from that in Schoen et al. (2013). The modification allows us to detect interesting strength-2 designs (for which $A_3 > 0$) with many estimable two-factor interaction effects, while the original W_2 criterion leads to the selection of strength-3 designs (for which $A_3 = 0$) with fewer estimable two-factor interaction effects. However, strength-3 designs will still be optimal in terms of the W_3 criterion, since any design for which $A_3 = 0$ also has a zero F_3 vector. In conclusion, by inserting $-R$ in the W_2 minimization vector, we increase the difference between the W_2 and W_3 criteria.

Besides $-R$, the elements of the W_2 criterion's minimization vector include the generalized length-3 word count A_3 of the two-level treatment factors, the treatment factors' F_3 vector, the block count $A_3^{(btot)}$, the generalized length-4 word count A_4 of the treatment design, and its F_4 vector. Positioning the F_3 vector immediately behind the A_3 value emphasizes our desire to avoid strong correlations between main effect contrast vectors and two-factor interaction contrast vectors among all design options with the same length-3 word count. For a similar reason, we position the F_4 vector behind the A_4 value. Cheng et al. (2004) considered separate criteria for generalized word counts A_i and for confounding frequency vectors F_i. We combine these elements to reduce the number of different criteria to consider.

In the W_2 criterion's minimization vector, the block count $A_3^{(btot)}$ is preceded by the generalized length-3 word count A_3 and itself precedes the length-4 count A_4. So, minimizing the confounding of the two-factor interactions with the blocks is prioritized over minimizing the aliasing among the two-factor interactions. This makes sense especially when large block effects are anticipated or when the number of blocks is small, in which cases the available inter-block information about confounded interaction effects is very limited and confounding of two-factor interactions with the blocks should be avoided to the largest possible extent.

Several authors also consider a W_1 criterion, which switches the positions of the block count $A_3^{(btot)}$ and the generalized length-4 word count A_4 (Cheng and Wu, 2002; Cheng et al., 2004; Schoen et al., 2013). We

Table 3: The W_2, W_2^- and W_3 criteria's minimization vectors for the three designs in Table 1

W_2 criterion:	$-R$	A_3	F_3	$A_3^{(btot)}$	A_4	F_4
Design 1	−21	0	(0,0,0)	3	14	(0,0,126)
Design 2	−24	1.11	(0,0,10)	3	10.89	(0,0,98)
Design 3	−21	0	(0,0,0)	7.5	14	(0,0,126)
W_2^- criterion:		A_3	F_3	$-A_3^{(btot)}$	A_4	F_4
Design 1		0	(0,0,0)	−3	14	(0,0,126)
Design 2		1.11	(0,0,10)	−3	10.89	(0,0,98)
Design 3		0	(0,0,0)	−7.5	14	(0,0,126)
W_3 criterion:			F_3			FA_3
Design 1			(0,0,0)			(0,0,18)
Design 2			(0,0,10)			(0,0,18)
Design 3			(0,0,0)			(0,9,18)

believe that this criterion is not useful for the cases with three of four blocks studied here, for the reasons that led us to recommend the W_2 criterion. Nevertheless, we did search for blocking arrangements that are optimal with respect to the W_1 criterion. It turned out that these blocking arrangements only differ from the W_2 optimal blocking arrangements for 7- and 8-factor 28-run designs in 7 blocks. We therefore only pay attention to the W_1 criterion for these two cases.

2.3.2 The W_2^- criterion

Schoen et al. (2013) introduced the W_2^- criterion to facilitate the detection of active main effects, by maximizing the confounding between the two-factor interaction effects and the blocks. In this way, the mean squared error of the main effects model would be inflated as little as possible due to any active two-factor interactions, resulting in the largest possible test statistic values for the treatment factors' main effects' significance tests.

Schoen et al. (2013) also discuss a W_1^- criterion, which reverses the prioritization of the maximization of the confounding between two-factor interactions and blocks and the minimization of the length-4 word count A_4. However, as the aim is to estimate a simple main-effects-only model, prioritizing the maximization of the confounding between two-factor interactions and blocks is more useful than minimizing the aliasing among two-factor interactions. For this reason, we do not use the W_1^- criterion in the present paper.

2.3.3 The W_3 criterion

The W_3 criterion is based on the F_3 and FA_3 vectors. Loosely speaking, it first minimizes the number of two-factor interactions that are severely confounded with the main effects and, subject to this, it minimizes the number of two-factor interactions that are substantially confounded with the blocks. It is, of course, possible to extend the minimization vector with the F_4 vector. However, for the designs considered in this paper, adding the F_4 vector to the minimization vector does not lead to different optimal designs.

2.4 Illustration of the three criteria

The designs 1, 2 and 3 in Table 1 are optimal according to the W_3 criterion, the W_2 criterion and the W_2^- criterion, respectively. The W_2^- optimal design involves the same treatments as the W_3 optimal design, but the treatments are assigned to the three blocks in a different fashion. We were able to establish the optimality of the three design options in terms of the three criteria by means of an exhaustive search over all nonisomorphic designs involving nine two-level (treatment) factors and one three-level (blocking) factor produced by the enumeration algorithm of Schoen et al. (2010).

For each of the three blocking arrangements in Table 1, Table 3 shows the W_2, W_2^- and W_3 criteria's minimization vectors. The second design option in Table 1 outperforms the other two designs in terms of the W_2 criterion, because it has the interaction model matrix with the largest rank, namely 24. The third design in Table 1 is best according to the W_2^- criterion, because it has the largest block count $A_3^{(btot)}$ of the

Table 4: Numbers of nonisomorphic orthogonally blocked two-level designs with 24 and 28 runs

Number of	24 runs			28 runs
factors	3 blocks	4 blocks	6 blocks	7 blocks
2	4	4	4	4
3	29	25	25	25
4	573	552	226	371
5	28,745	21,757	2,663	21,502
6	1,089,168	457,768	19,323	395,598
7	14,576,216	3,113,669	58,112	1,422,094
8	57,436,095	8,168,256	65,679	1,005,490
9	71,157,023	12,605,571	26,454	135,569
10	33,893,515	15,119,461	12,243	4,296
11	10,266,252	14,961,206	4,882	104
12	3,305,030	12,096,092	1,543	21
13	981,180	7,855,020	277	
14	220,993	4,066,838	45	
15	32,567	1,665,918		
16	2,282	532,484		
17		129,122		
18		22,880		
19		2,758		
20		238		

three designs (7.5), in addition to a zero A_3 value and a zero F_3 vector. Finally, the first design option in Table 1 outperforms the second design option in terms of the W_3 criterion, because it has a better F_3 vector. It outperforms the third design option in terms of the W_3 criterion, because it has the same ideal F_3 vector and a better FA_3-vector.

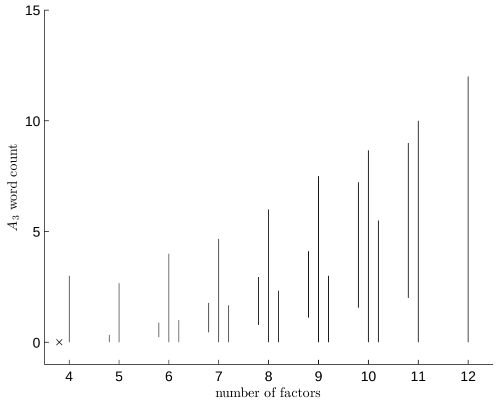
A preference for one of the three designs depends on the exact goal of the experimenters. If interactions are considered as nuisance effects, designs 1 and 3 both stand out. Design 3 is preferred if one wants a main effects model with standard errors based on the mean squared error, while there might be a number of small two-factor interactions. The mean squared error is less biased in this design than in the other two designs. Design 1 is preferred if one wishes to remove one or two more substantial interactions from the mean squared error. Finally, design 2 is recommended if finding interaction effects is the main purpose of the experiment.

3 Results

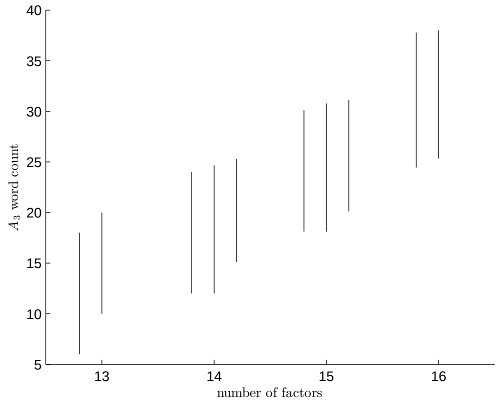
Table 4 shows the numbers of nonisomorphic orthogonal 24-run two-level treatment designs involving a three-level, a four-level and a six-level blocking factor as well as the numbers of non-isomorphic orthogonal 28-run two-level treatment designs involving a seven-level blocking factor (Schoen et al., 2010). By an exhaustive evaluation of all these designs in terms of the three criteria, we were able to identify all optimal orthogonally blocked two-level designs involving 24 or 28 runs. Our exploration of the complete catalog was performed in the same way as the search done by Schoen et al. (2010) for smaller designs. As the number of nonisomorphic designs is very large, this exploration was computationally demanding. The appendix to this paper includes tables showing the F_3, F_4 and FA_3 vectors along with the ranks of the interaction model matrices for the optimal blocking arrangements we identified. Tables containing the optimal blocking arrangements themselves are available from the authors.

We visualize the generalized length-3 word count and the block count of the optimal designs involving four or more two-level treatment factors in Figure 1. Panels (a) and (b) of the figure show results for 24-run designs involving three blocks. Panels (c) and (d) show results for 24-run designs involving four blocks. Panel (e) visualizes the results for 24-run designs involving six blocks, and, finally, panel (f) contains results for 28-run designs and seven blocks.

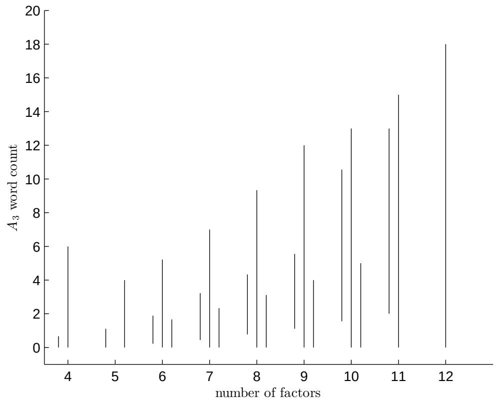
Every vertical line in the panels corresponds to one optimal blocking pattern we identified. The lower



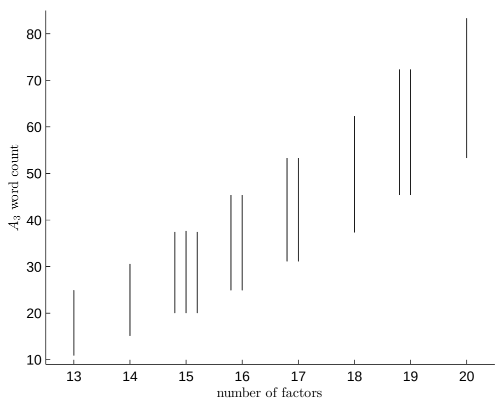
(a) 24 runs, 3 blocks, 4–12 factors



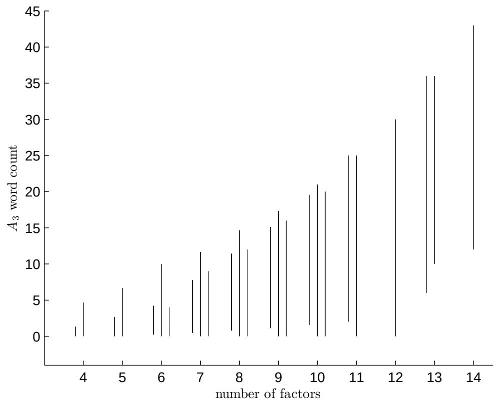
(b) 24 runs, 3 blocks, 13–16 factors



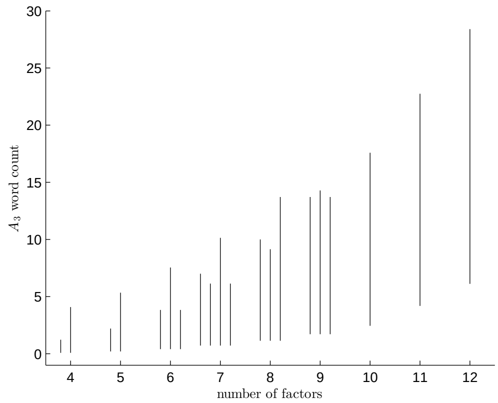
(c) 24 runs, 4 blocks, 4–12 factors



(d) 24 runs, 4 blocks, 13–20 factors



(e) 24 runs, 6 blocks, 4–14 factors



(f) 28 runs, 7 blocks, 4–12 factors

Figure 1: Properties of the optimal blocking arrangements of 24-run and 28-run orthogonal designs. The lower ends of the vertical lines indicate the A_3 values for the treatment designs, while the lines' lengths represent the block counts.

end of each line indicates the generalized length-3 word count A_3 of the two-level treatment design, while the length of the line indicates the block count $A_3^{(btot)}$. As a result, a low lower end of a line means a small degree of aliasing between the treatment factors' main effects and their two-factor interactions. A short vertical line means a small degree of confounding between the treatment factors' two-factor interaction effects and the blocks. A high higher end of a vertical line means a severe aliasing of the two-factor interactions with the main effects and/or confounding of the two-factor interactions with the block effects. So, the higher end of each line provides a measure of the total confounding and aliasing of the two-factor interactions, $A_3 + A_3^{(btot)}$.

Generally, the vertical lines appear in clusters of three and correspond, from left to right, to the W_2 optimal, W_2^- optimal, and W_3 optimal designs for the number of treatment factors displayed on the horizontal axis. Whenever a cluster contains only one or two designs, this means that we identified designs that are optimal with respect to more than one criterion.

3.1 24-run designs in three blocks

Panel (a) of Figure 1 shows the results for 24-run designs involving 4–12 treatment factors and three blocks. For each number of treatment factors, the plot shows one or two designs of strength 3, which have a generalized length-3 word count, A_3, of zero. These designs are optimal in terms of the W_2^- and W_3 criteria, because these criteria prioritize the minimization of the A_3 value. In all of these cases, the same two-level treatment design is used in the W_2^- and W_3 optimal blocked designs. For the case of nine treatment factors, that treatment design, along with its W_2^- and W_3 optimal blocking arrangements, is shown in Table 1. Design 3 in Table 1 corresponds to the W_2^- optimal design, while design 1 is the W_3 optimal design.

All designs involving four or five treatment factors have a strength-3 treatment design. The small cross corresponds to a four-factor design that does not involve any confounding of the two-factor interactions with the blocks. It consists of two replicates of the regular 2^{4-1}_{IV} design and one replicate of that design's foldover. Each of the three replicates is assigned to a different block. The design is W_2 optimal and W_3 optimal. The second design with 4 factors involves three replicates of the regular 2^{4-1}_{IV} design, each assigned to a different block. This design is W_2^- optimal. There are two designs with five treatment factors. The first one is used in the W_2 optimal and W_3 optimal blocking arrangement, while the second is used in the W_2^- optimal blocking arrangement. Unlike the designs with four treatment factors, the five-factor treatment designs cannot be constructed using regular fractions.

The W_2 optimal designs involving 6–11 treatment factors have a nonzero generalized length-3 word count A_3, but, nevertheless, a larger value for the rank criterion than the alternative designs with the same number of factors. The nonzero A_3 value of the W_2 optimal designs can be recognized in the plot by looking at the lower end of the leftmost vertical line for each number of factors in panel (a): that line does not start at zero. The longest vertical line for each number of factors in the plot always corresponds to the W_2^- optimal design. This is because that design maximizes the confounding of the two-factor interaction effects with the blocks, as opposed to the W_2 and W_3 optimal designs. Another pattern that is visible in the figure is that the upper ends of the vertical lines increase with the number of factors. This shows that, for a given number of experimental runs and a given criterion, studying larger numbers of factors results in larger degrees of aliasing of two-factor interactions and/or confounding of the two-factor interactions with the blocks. Finally, for the case of 11 factors, there exists one design that simultaneously optimizes the W_2^- and W_3 criteria. For the case of 12 treatment factors, there is a single blocking arrangement that is optimal in terms of all three criteria. The treatment design in that blocking arrangement is a folded-over 12-run Plackett-Burman design (Plackett and Burman, 1946). Each block in the blocking arrangement consists of four mirror image pairs.

Panel (b) of Figure 1 shows the results for 24-run designs involving 13–16 factors and three blocks. For these numbers of factors, no strength-3 designs exist and the rank of the interaction model matrix equals 24 for all designs under study. Therefore, for these numbers of factors, the W_2 and W_2^- criteria are more alike than for smaller numbers of factors, where the rank for strength-3 designs is always lower than that for the best strength-2 designs. More specifically, for 13–16 factors, the W_2 and W_2^- criteria favor the same two-level treatment designs. For the case of 13 factors, the two criteria even lead to the same W_2 and W_2^- optimal blocking arrangement with a maximum absolute type-3 correlation of 2/3. In contrast, the W_3 optimal blocking arrangement for 13 factors has a treatment design with a maximum absolute type-3 correlation of 1/3 only and a smaller block count $A_3^{(btot)}$. While these are attractive properties, the total generalized length-3 word count of the W_3 optimal blocking arrangement, $A_3 + A_3^{(btot)}$, is higher than that of the W_2 and W_2^- optimal blocking arrangement. So, neither of the two designs outperforms the other in all possible ways.

For 14 and 15 factors, the W_2 and W_2^- optimal blocking arrangements use the same treatment design, but the assignment of the treatments to the blocks differs. The difference between these arrangements is too small to be of practical importance. It is barely visible in the plot. The comparison between the W_2 and W_2^- optimal designs, on the one hand, and the W_3 optimal designs, on the other hand, is analogous to that for the 13-factor designs.

Finally, the selected 16-factor blocking arrangements are either W_2 and W_2^- optimal or W_3 optimal. Both have a two-level treatment design with absolute type-3 correlations of 2/3. There is very little practical difference between these blocking arrangements.

3.2 24-run designs in four blocks

Panel (c) of Figure 1 shows the results for 24-run treatment designs involving 4–12 treatment factors and four blocks. The salient features of these designs strongly resemble those of the designs arranged in three blocks. For example, the blocking arrangements that are optimal in terms of the W_2^- and W_3 criteria also have a strength-3 treatment design (and thus a zero A_3 value). The plot in panel (c) also shows that, all other things being equal, it is harder to achieve good blocking arrangements for four blocks of size 6 than for three blocks of size 8. As a matter of fact, the upper ends of the vertical lines in the figure's panel (c) are higher than those in panel (a). In other words, the block count, measuring the confounding between the two-factor interactions and the blocks, is larger.

Panel (d) of Figure 1 shows the results for 24-run designs involving 13–20 treatment factors and arranged in four blocks. The most striking feature of the plot is that, for each given number of factors, the optimal blocking arrangements have the same generalized length-3 word counts A_3, and, except for the case of 15 factors, the same block counts $A_3^{(btot)}$. Indeed, except for 15 factors, the optimal blocking arrangements we found only differ in terms of the F_4 and FA_3 vectors. The 15-factor optimal blocking arrangements also differ in terms of the $A_3^{(btot)}$ value (but the small difference in $A_3^{(btot)}$ value is hardly visible in the plot).

Table 5 shows the F_3, F_4 and FA_3 vectors and the block counts $A_3^{(btot)}$ for the W_2, W_2^- and W_3 optimal arrangements of 15-factor 24-run designs in four blocks. The rank of the interaction is 24 for each of these 15-factor designs. All the designs tabulated possess the smallest A_3 value possible for a 15-factor design in 24 runs, namely 20. They further have the same A_4 value of 67.67.

Among all designs with the most attractive A_3 value and F_3 vector, $F_3(24, 16, 8) = (0, 0, 180)$, the W_2 criterion first minimizes the block count $A_3^{(btot)}$ and the A_4 value, and then sequentially minimizes the elements of the F_4 vector. In contrast, among all designs with the most attractive A_3 value and F_3 vector, the W_2^- criterion first maximizes the block count $A_3^{(btot)}$ and then sequentially minimizes the A_4 value and the elements of the F_4 vector. The minimum value for the block count $A_3^{(btot)}$ among all designs with the most attractive A_3 value and F_3 vector is 157/9, while the maximum value is 159/9. This small difference causes the vertical lines for the W_2 and W_2^- optimal blocking arrangements in the figure's panel (d) to be nearly equally long. Finally, among all designs with the most attractive F_3 vector, the W_3 criterion sequentially minimizes the FA_3 vector. All these features are listed in Table 5. What cannot be seen from the table is that the three tabulated designs result from an exhaustive search.

In case 16, 17 and 19 factors are studied, the W_2 and W_2^- optimal blocking arrangements are identical, but they differ from the W_3 optimal blocking arrangement. For 13, 14, 18 and 20 factors, the optimal

Table 5: Characteristics of the W_2, W_2^- and W_3 optimal blocking arrangements of 15-factor 24-run treatment designs in four blocks. The elements of the FA_3 vector correspond to length-3 word counts of 1, 5/9, 3/9 and 1/9.

$F_3(24, 16, 8)$			$F_4(24, 16, 8)$			FA_3				$A_3^{(btot)}$	Optimality
0	0	180	2	21	507	1	7	8	89	157/9	W_2
0	0	180	5	12	516	2	5	9	89	159/9	W_2^-
0	0	180	2	23	499	0	9	8	88	157/9	W_3

blocking arrangements all coincide.

3.3 24-run designs in six blocks

Panel (e) of Figure 1 shows the results for 24-run designs involving 4–14 treatment factors and arranged in six blocks. Because of the larger number of blocks, there is more confounding between the two-factor interaction effects and the blocks. Again, the blocking arrangements that are optimal in terms of the W_2^- and W_3 criteria have a treatment design with a zero A_3 value when 12 or fewer treatment factors are studied.

The largest number of two-level treatment factors that can be handled with 24 runs in six blocks is 14, whereas 16 factors can be studied when there are only three blocks and 20 factors can be studied when there are four blocks. Studying more than 14 treatment factors is possible using a six-block 24-run design, but only with a design that is not orthogonally blocked.

3.4 28-run designs in seven blocks

The last series of blocking arrangements we discuss involves 28 runs and seven blocks. Panel (f) of Figure 1 shows the characteristics of the optimal blocking arrangements we identified for that case. The plot shows two different optimal blocking arrangements for four and five factors, three different blocking arrangements for six, eight and nine factors, four different blocking arrangement for seven factors and a single blocking arrangement for 10–12 factors.

The blocking arrangements for seven and eight factors deserve special attention, because these are the only instances we encountered where the W_2 optimal designs differ from the W_1 optimal designs. Recall that, in the W_2 criterion's minimization vector, the block count $A_3^{(btot)}$ precedes the generalized length-4 word count A_4, while, in the W_1 criterion's minimization vector, the generalized length-4 word count A_4 precedes the block count $A_3^{(btot)}$.

Table 6 shows the rank R of the interaction model matrix, the generalized length-4 word counts A_4 and the block counts $A_3^{(btot)}$ for the W_1, W_2, W_2^- and W_3 optimal 7-factor and 8-factor 28-run designs arranged in seven blocks. All four 7-factor designs have the same F_3 vector. Therefore, the W_1 criterion favors the design with the smallest A_4 value, while the W_2 criterion favors the design with smallest $A_3^{(btot)}$ value, the W_2^- criterion favors the design with the largest $A_3^{(btot)}$ value and ignores the rank, and the W_3 criterion favors the design with best FA_3 vector (ignoring the A_4 value).

For the 8-factor designs, the results are similar to those for the 7-factor designs, although, for that case, the W_2 and W_3 criteria produce the same optimal blocking arrangement.

Finally, for nine factors, the W_2 optimal design has the same generalized length-3 word count A_3 and the same block count $A_3^{(btot)}$ as the W_3 optimal design. The designs differ in the rank of the interaction model matrix, the A_4 value, the F_4 vector and the FA_3 vector.

It is natural to wonder what benefits the seventh block of size 4 in a 28-run design offers when compared to a 24-run design arranged in six blocks of size 4. As discussed in Section 2, all type-3 correlations for 28-run designs are larger than zero. This implies that the main effects are always partially aliased with the two-factor interactions when a 28-run design is used. So, all 28-run designs arranged in seven blocks have a larger generalized length-3 word count A_3 than the corresponding 24-run design options involving six

Table 6: Characteristics of the W_1, W_2, W_2^- and W_3 optimal arrangements for 28-run 7- and 8-factor designs in seven blocks. The elements of the FA_3 vector correspond to length-3 word counts of 6/7, 4/7 and 2/7.

# factors	R	A_4	$A_3^{(btot)}$	FA_3			Optimality
7	28	1.04	6.29	0	5	12	W_1
7	28	1.37	5.43	0	3	13	W_2
7	26	2.18	9.43	3	6	12	W_2^-
7	28	1.53	5.43	0	1	17	W_3
8	28	3.06	8.86	0	8	15	W_1
8	28	3.71	8.00	0	4	20	W_2, W_3
8	27	4.36	12.57	4	8	16	W_2^-

blocks. However, the additional runs in the 28-run designs also result in a smaller amount of confounding of the two-factor interactions with the blocks: all 28-run design options we recommend have a smaller block count $A_3^{(btot)}$ than the corresponding 24-run design options. Therefore, if the emphasis in an experiment is on detecting main effects, we recommend the 24-run W_2^- optimal designs, despite the small increase in standard error by a factor of 1.08. If one anticipates a few two-factor interactions along with the main effects, we prefer the 28-run W_2 or W_3 optimal designs.

3.5 Comparison with the best designs for complete randomization

Most of the treatment designs that appear in the blocking arrangements that are optimal in terms of the three criteria we study here are also optimal or nearly optimal when run in a completely randomized fashion. Specifically, most of these designs are also best, second best or third best when all possible treatment designs are ordered according to the F_3 and F_4 vectors or according to the (A_3, F_3, A_4, F_4) vector (Schoen et al., 2015). Several top-3 treatment designs, however, cannot be orthogonally blocked when three, four, six or seven blocks are desired:

1. Three blocks, 24 runs, 14–16 factors: all top-3 treatment designs according to the F_3 and F_4 vectors.
2. Four blocks, 24 runs, 13, 14, 17 and 19 factors: all top-3 treatment designs according to the F_3 and F_4 vectors.
3. Four blocks, 24 runs, 15 and 16 factors: all top-3 treatment designs with the smallest (A_3, F_3, A_4, F_4) vector.
4. Six blocks, 24 runs, 14 factors: all top-3 treatment designs according to the F_3 and F_4 vectors.
5. Seven blocks, 28 runs, 7 and 8 factors: the best treatment design according to the F_3 and F_4 vectors.
6. Seven blocks, 28 runs, 9–12 factors: all top-3 treatment designs according to the F_3 and F_4 vectors.

4 Discussion

This paper makes a large number of orthogonal blocking arrangements available for orthogonal two-level designs involving 24 and 28 runs. To identify optimal blocking arrangements, we explored complete catalogs of orthogonal designs with several two-level treatment factors and one multi-level blocking factor. We focused on three optimality criteria. We restricted attention to orthogonal designs in orthogonal blocks, because these designs result in maximum precision for the main effect estimates when estimating a model including the block effects and the treatment factors' main effects only. Depending on the specific case, when using the designs we present for estimating a main effects model, the main effect estimates are either not at all biased when a few interactions are active, or only to a small extent.

For some of the cases considered here, the restriction to orthogonal blocking comes at a price. Table 4 showed that orthogonal blocking arrangements of the 24-run orthogonal designs involving more than 16 two-level factors do not exist if three blocks are desired. Similarly, orthogonal blocking arrangements of the 24-run orthogonal designs involving more than 14 two-level factors do not exist if six blocks are desired, and orthogonal blocking arrangements of the 28-run orthogonal designs involving more than 12 two-level factors do not exist for any number of blocks. Further, a comparison of the results presented here with those in Schoen et al. (2015) for two-level designs with 24 or 28 runs in the absence of blocking revealed a few cases where orthogonal blocking was only feasible for suboptimal treatment designs.

It is, of course, also possible to create blocking arrangements that are not orthogonally blocked, using principles from optimal experimental design. We would welcome follow-up research investigating the capabilities of non-orthogonal blocking of good treatment designs with 24 and 28 runs. The most important characteristic of a screening design is its potential to detect active main effects and two-factor interactions. A simulation tool to study the power for detecting active main effects and two-factor interactions in blocked experiments would provide an objective basis for comparing orthogonally blocked (W_2, W_2^- and W_3 optimal) and nonorthogonally blocked designs.

Acknowledgement

The research of the first author was supported by the Flemish Fund for Scientific Research FWO.

References

Angelopoulos, P., Evangalaras, H., Koukouvinos, C., and Lappas, E. (2007). An effective step-down algorithm for the construction and the identification of nonisomorphic orthogonal arrays. *Metrika*, 66:139–149.

Bisgaard, S. (1994). A note on the definition of resolution for blocked 2^{k-p} designs. *Technometrics*, 36:308–311.

Bulutoglu, D. A. and Ryan, K. J. (2015). Algorithms for finding generalized minimum aberration designs. *Journal of Complexity*, 31:577–589.

Cheng, C. S., Mee, R. W., and Yee, O. (2008). Second order saturated orthogonal arrays of strength three. *Statistica Sinica*, 18:105–119.

Cheng, S. W., Li, W., and Ye, K. Q. (2004). Blocked nonregular two-level factorial designs. *Technometrics*, 46:269–279.

Cheng, S. W. and Wu, C. F. J. (2002). Choice of optimal blocking schemes in two-level and three-level designs. *Technometrics*, 44:269–277.

Deng, L. Y. and Tang, B. (1999). Generalized resolution and minimum aberration criteria for Plackett-Burman and other nonregular factorial designs. *Statistica Sinica*, 9:1071–1082.

Deng, L. Y. and Tang, B. (2002). Design selection and classification for Hadamard matrices using generalized minimum aberration criteria. *Technometrics*, 44:173–184.

Evangelaras, H., Koukouvinos, C., and Lappas, E. (2007). Further contributions to nonisomorphic two level orthogonal arrays. *Journal of Statistical Planning and Inference*, 137:2080–2086.

Ingram, D. and Tang, B. (2005). Minimum G aberration design construction and design tables for 24 runs. *Journal of Quality Technology*, 37:101–114.

Li, Y., Deng, L. Y., and Tang, B. (2004). Design catalog based on minimum G-aberration. *Journal of Statistical Planning and Inference*, 124:219–230.

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

Montgomery, D. C. (2009). *Design and analysis of experiments.* 7th edition, Wiley, New York.

National Bureau of Standards (1957). *Fractional factorial designs for factors at two levels.* Applied Mathematics Series 48; Washington DC: US Government Printing Office.

Plackett, R. L. and Burman, J. P. (1946). The design of optimum multifactorial experiments. *Biometrika*, 33:305–325.

Rao, C. R. (1947). Factorial experiments derivable from combinatorial arrangements of arrays. *Journal of the Royal Statistical Society Supplement*, 9:128–139.

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., Sartono, B., and Goos, P. (2013). Optimum blocking for general resolution-3 designs. *Journal of Quality Technology*, 45:166–187.

Schoen, E. D., Vo-Thanh, N., and Goos, P. (2015). 24- and 28-run two-level orthogonal designs. Working paper 2015-016, University of Antwerp, Faculty of Applied Economics.

Sitter, R. R., Chen, J., and Feder, M. (1997). Fractional resolution and minimum aberration in blocked 2^{n-k} designs. *Technometrics*, 39:382–390.

Sun, D. X., Li, W., and Ye, K. Q. (2008). An algorithm for sequentially constructing nonisomorphic orthogonal designs and its applications. *Statistics and Applications*, 6:144–158.

Sun, D. X., Wu, C. F. J., and Chen, Y. (1997). Optimal blocking schemes for 2^n and 2^{n-p} designs. *Technometrics*, 39:298–307.

Tang, B. and Deng, L. Y. (1999). Minimum G_2-aberration for nonregular fractional factorial designs. *Annals of Statistics*, 27:1914–1926.

Wu, C. F. J. and Hamada, M. S. (2009). *Experiments: Planning, Analysis, and Parameter Design Optimization*. 2nd edition, Wiley, New York, NY, USA.

Xu, H. (2006). Blocked regular fractional factorial designs with minimum aberration. *Annals of Statistics*, 34:2534–2553.

Xu, H. and Deng, L. Y. (2005). Moment aberration projection for nonregular fractional factorial designs. *Technometrics*, 47:121–131.

Xu, H. and Wu, C. F. J. (2001). Generalized minimum aberration for asymmetrical fractional factorial designs. *Annals of Statistics*, 29:1066–1077.

Appendix

Table A1: Properties of optimal arrangements of 24-run designs in three blocks. The elements of the FA_3 vector correspond to length-3 word counts of 2/3, 1/2 and 1/6.

ID	$F_3(16,8)$		$F_4(24,16,8)$			FA_3			Rank	Optimality
4.1	0	0	0	0	1	0	0	0	11	W_2, W_3
4.2	0	0	1	0	0	0	6	0	8	W_2^-
5.1	0	0	0	0	5	0	0	2	16	W_2, W_3
5.2	0	0	0	0	5	1	2	6	16	W_2^-
6.1	0	2	0	0	9	0	0	4	22	W_2
6.2	0	0	0	0	15	3	0	12	18	W_2^-
6.3	0	0	0	0	15	0	0	6	18	W_3
7.1	0	4	0	0	21	0	0	8	24	W_2
7.2	0	0	0	0	35	2	3	11	19	W_2^-
7.3	0	0	0	0	35	0	0	10	19	W_3
8.1	0	7	0	0	53	0	0	13	24	W_2
8.2	0	0	0	0	70	2	4	16	20	W_2^-
8.3	0	0	0	0	70	0	0	14	20	W_3
9.1	0	10	0	0	98	0	0	18	24	W_2
9.2	0	0	0	0	126	0	9	18	21	W_2^-
9.3	0	0	0	0	126	0	0	18	21	W_3
10.1	0	14	0	0	168	1	2	24	24	W_2
10.2	0	0	0	0	210	2	4	32	22	W_2^-
10.3	0	0	0	0	210	0	3	24	22	W_3
11.1	0	18	0	0	270	1	3	29	24	W_2
11.2	0	0	0	0	330	3	5	33	23	W_2^-
11.3	0	0	0	0	330	0	10	30	23	W_3
12.1	0	0	0	0	495	0	12	36	24	W_2, W_2^-, W_3
13.1	6	30	0	0	495	3	6	42	24	W_2, W_2^-
13.2	0	90	1	0	366	0	3	51	24	W_3
14.1	4	92	1	0	540	1	4	56	24	W_2
14.2	4	92	1	0	540	4	4	48	24	W_2^-
14.3	0	136	3	10	402	0	0	61	24	W_3
15.1	3	151	3	0	630	0	6	54	24	W_2
15.2	3	151	3	0	630	1	3	63	24	W_2^-
15.3	0	181	3	4	578	0	1	63	24	W_3
16.1	4	204	6	0	766	1	2	70	24	W_2, W_2^-
16.2	3	216	3	14	721	0	0	76	24	W_3

Table A2: Properties of optimal arrangements of 24-run designs in four blocks. The elements of the FA_3 vector correspond to length-3 word counts of 1, 5/9, 1/3 and 1/9.

ID	$F_3(8)$	$F_4(24,16,8)$			FA_3				Rank	Optimality
4.1	0	0	0	1	0	0	0	6	11	W_2, W_3
4.2	0	1	0	0	6	0	0	0	8	W_2^-
5.1	0	0	0	5	0	0	0	10	16	W_2, W_3
5.2	0	0	0	5	3	0	1	6	16	W_2^-
6.1	2	0	0	9	0	0	0	15	22	W_2
6.2	0	0	0	15	1	5	2	7	18	W_2^-
6.3	0	0	0	15	0	0	0	15	18	W_3
7.1	4	0	0	21	0	0	2	19	24	W_2
7.2	0	0	0	35	0	6	9	6	19	W_2^-
7.3	0	0	0	35	0	0	0	21	19	W_3
8.1	7	0	0	53	0	1	0	27	24	W_2
8.2	0	0	0	70	0	8	12	8	20	W_2^-
8.3	0	0	0	70	0	0	0	28	20	W_3
9.1	10	0	0	98	0	0	2	34	24	W_2
9.2	0	0	0	126	0	0	36	0	21	W_2^-
9.3	0	0	0	126	0	0	0	36	21	W_3
10.1	14	0	0	168	1	0	14	30	24	W_2
10.2	0	0	0	210	0	0	36	9	22	W_2^-
10.3	0	0	0	210	0	0	0	45	22	W_3
11.1	18	0	0	270	1	0	18	36	24	W_2
11.2	0	0	0	330	3	0	28	24	23	W_2^-
11.3	0	0	0	330	0	10	20	25	23	W_3
12.1	0	0	0	495	0	18	12	36	24	W_2, W_2^-, W_3
13.1	98	1	4	342	0	8	8	62	24	W_2, W_2^-, W_3
14.1	136	2	12	403	0	4	16	71	24	W_2, W_2^-, W_3
15.1	180	2	21	507	1	7	8	89	24	W_2
15.2	180	5	12	516	2	5	9	89	24	W_2^-
15.3	180	2	23	499	0	9	8	88	24	W_3
16.1	224	4	32	648	0	8	16	96	24	W_2, W_2^-
16.2	224	12	16	640	0	0	32	88	24	W_3
17.1	280	2	83	686	0	16	0	120	24	W_2, W_2^-
17.2	280	12	16	864	0	0	32	104	24	W_3
18.1	336	9	81	927	0	9	18	126	24	W_2, W_2^-, W_3
19.1	408	4	38	1480	0	11	14	146	24	W_2, W_2^-
19.2	408	9	81	1263	0	9	18	144	24	W_3
20.1	480	5	48	1848	0	12	16	162	24	W_2, W_2^-, W_3

Table A3: Properties of optimal arrangements of 24-run designs in six blocks. The elements of the FA_3 vector correspond to length-3 word counts of 1, 2/3, and 1/3.

ID	$F_3(16,8)$		$F_4(24,8)$		FA_3			Rank	Optimality
4.1	0	0	0	1	0	0	4	11	W_2, W_3
4.2	0	0	0	1	2	4	0	11	W_2^-
5.1	0	0	0	5	0	1	6	16	W_2, W_3
5.2	0	0	0	5	0	10	0	16	W_2^-
6.1	0	2	0	9	0	2	8	22	W_2
6.2	0	0	0	15	0	15	0	18	W_2^-
6.3	0	0	0	15	0	2	8	17	W_3
7.1	0	4	0	21	0	4	14	24	W_2
7.2	0	0	0	35	0	15	5	19	W_2^-
7.3	0	0	0	35	0	6	15	19	W_3
8.1	0	7	0	53	1	5	19	24	W_2
8.2	0	0	0	70	4	8	16	20	W_2^-
8.3	0	0	0	70	0	8	20	20	W_3
9.1	0	10	0	98	0	9	24	24	W_2
9.2	0	0	0	126	1	14	21	21	W_2^-
9.3	0	0	0	126	0	12	24	21	W_3
10.1	0	14	0	168	0	12	30	24	W_2
10.2	0	0	0	210	1	16	28	22	W_2^-
10.3	0	0	0	210	0	16	28	22	W_3
11.1	0	18	0	270	1	16	34	24	W_2
11.2	0	0	0	330	2	20	29	23	W_2^-
11.3	0	0	0	330	0	25	25	23	W_3
12.1	0	0	0	495	0	30	30	24	W_2, W_2^-, W_3
13.1	6	30	0	495	2	24	36	24	W_2, W_2^-
13.2	0	90	1	366	2	12	48	24	W_3
14.1	4	92	1	540	3	16	52	24	W_2, W_2^-, W_3

Table A4: Properties of optimal arrangements of 28-run designs in seven blocks. The elements of the FA_3 vector correspond to length-3 word counts of 6/7, 4/7 and 2/7.

ID	$F_3(12,4)$		$F_4(20,12,4)$			FA_3			Rank	Optimality
4.1	0	4	0	0	1	0	0	4	11	W_1, W_2, W_3
4.2	0	4	0	1	0	2	4	0	11	W_2^-
5.1	0	10	0	0	5	0	0	7	16	W_1, W_2, W_3
5.2	0	10	0	2	3	1	6	3	16	W_2^-
6.1	0	20	0	0	15	0	1	10	22	W_1, W_2
6.2	0	20	1	4	10	3	4	8	20	W_2^-
6.3	0	20	0	0	15	0	0	12	22	W_3
7.1	0	35	0	2	33	0	5	12	28	W_1
7.2	0	35	0	4	31	0	3	13	28	W_2
7.3	0	35	3	0	32	3	6	12	26	W_2^-
7.4	0	35	1	2	32	0	1	17	28	W_3
8.1	0	56	0	10	60	0	8	15	28	W_1
8.2	0	56	2	8	60	0	4	20	28	W_2, W_3
8.3	0	56	6	0	64	4	8	16	27	W_2^-
9.1	0	84	0	25	101	0	10	22	28	W_1, W_2
9.2	0	84	6	8	112	0	12	20	28	W_2^-
9.3	0	84	0	36	90	0	6	30	26	W_3
10.1	0	120	10	16	184	0	8	37	28	W_1, W_2, W_2^-, W_3
11.1	5	160	10	32	288	0	10	45	28	W_1, W_2, W_2^-, W_3
12.1	10	210	15	48	432	0	12	54	28	W_1, W_2, W_2^-, W_3