# Blocking versus Robustness in Industrial Contexts

Pere Grima, Lluís Marco-Almagro, Xavier Tort-Martorell
Department of Statistics and Operational Research
Universitat Politècnica de Catalunya – BarcelonaTech, Barcelona (Spain)

The paper discusses the similarities and differences between blocking factors (blocked designs) and noise factors (robust designs) in industrial two level factorial experiments. The discussion covers from the objectives of both design types and the nature of blocking and noise factors to the types of designs and the assumptions needed in each case. The conclusions are: the nature and characteristics of noise and blocking factors are equal or very similar; the designs used in both situations are also similar; and the main differences lie in the assumptions and the objectives. The paper argues that the objectives are not in conflict and can easily be harmonized. In consequence, we argue in favor of a unified approach that would clarify the issue, especially for students and practitioners.

## 1. Introduction

The discipline of design of experiments is often taught or presented as a unified theory without taking into account differences that depend on the science in which it is applied. For example, Box[1] and, more recently, Vining[2] show an important difference between agriculture and industry when designing experiments, specifically in regard to the time it takes to run an experiment. The duration is much shorter in industry – from a few minutes or hours up to, in the most extreme and rare cases, some days – while, in agriculture, it can easily last one or more years. This difference leads to designing larger experiments for agriculture while industry allows the sequential nature of experimentation to be used at its full potential.

There are obviously many common aspects, while relevant differences are usually treated in the texts concerning the science in question. One common aspect that is always emphasized is the importance of blocking and its importance in offsetting differences in experimental conditions.

It is well known that the original idea comes from Fisher, who faced problems arising from the lack of homogeneity in farm plots (hence the name "blocks"). The phrase in Box, Hunter and Hunter[3] (p. 93): "Block what you can and randomize what you cannot" has become classic. Virtually all books on design of experiments refer to it, regardless of their focus and level. In this article, we analyze the relationship between blocking and robust parameter design, and we do so from the point of view of experimentation in industry while concentrating on $2^k$ designs. We do not discuss whether the ideas raised here may also be useful in other areas, although this may very well be the case to at least some extent.

In the first part of this article, we review how blocking and robustness are presented in the three books on design of experiments that are probably the most widespread in industrial environments: Box, Hunter and Hunter[3]; Montgomery[4]; and Wu and Hamada[5]. Sections 2 and 3 present the objectives, design types and assumptions for block and robust designs, respectively. In Section 4, we argue in favor of unifying the objectives and types of designs, and, finally, we summarize the conclusions and further discuss the advantages of this unified approach.

## 2. Blocking

Regardless of the sector towards which they are oriented, all texts on design of experiments emphasize the importance of blocking, and they frequently introduce the concept via an example from agriculture. They explain that blocking factors are not the point of interest and that the objective is to eliminate the variability that these factors generate in order to improve precision when comparing treatments, which is the ultimate aim of the experiment.

In the industrial context, some typical examples of "blocking factors" are: raw material from different batches or suppliers; different operators; different machines; or a lack of homogenous conditions due to time elapsed between runs (which is often related to changes in environmental conditions, such as humidity and temperature). While in other contexts it may be true that the blocking factors are of no interest, we should consider that this is not so obvious in industry and thus ask ourselves: Are we truly not interested in knowing if and how much the response of interest is affected by changes in raw materials or environmental conditions? Perhaps this is one reason behind the difficulty in finding in the literature practical industrial applications for blocking, despite its theoretical importance and usefulness. Four relatively recent papers have compiled a number of case studies on design of experiments: Bisgaard[6]; Prvan and Street[7]; Itzarbe et al.[8]; and Sharif et al.[9] Among these four papers, only 9 of the 425 references deal with blocking factors. What is more, they do so only in the context of relatively complex

designs or response surface methodology studies. Additionally, and focusing on $2^k$ and $2^{k-p}$ designs, it is worth noting that the three books on design of experiments that are most widespread and referenced in the industrial sector, and which often illustrate the concepts with real cases either have no blocking examples (Wu and Hamada[5]), the examples do not come from industry (Box et al.[3]) or that they are not real cases (Montgomery[4]) (see Table 1). The exception to the lack of real cases occurs when blocking is a consequence of sequentially adding experiments (as in Sanders et al.[10]), especially when using response surface methodology (as in Jensen and Kowalski[11]).

Table 1: Examples of blocked $2^k$ designs in influential industrial DOE books

| Book | Context | Blocking factors | Reference |
|---|---|---|---|
| Box, Hunter and Hunter[3] | Time for traveling a route on bike | Tiring and getting in shape | Not industrial (student project) |
| Montgomery[4] | Target detection on a radar | Operator | Not cited |
| | Chemical production process (example made up by modifying a real randomized experiment) | Raw material batch | Not cited |
| | Developing a plasma etching process | Shift | Not cited |
| | Turbine engine production process | Spindle | Not cited |
| Wu and Hamada[5] (in the exercises) | Door panel stamping | Day | Not cited |
| | Effort to close the hatch in a new car design | Car | Young[12] |

As is well known, the entire blocking theory is based on the assumption that block-by-treatment interactions are negligible. We have been unable to find any justification, neither theoretical nor empirical, for this generic assertion beyond general statements of the type: "the validity of the assumption has been confirmed in many empirical studies", which appears in Wu and Hamada[5] (p. 129). Naturally, this assumption, like all others, should be checked in each particular case. Furthermore, as it could not be otherwise, many authors recommend doing so. For example, Box, Hunter and Hunter[3] discuss the importance of checking the additivity hypothesis (p. 151) by analyzing the residuals and indicating if in some cases a transformation may correct problems due to lack of additivity; while Wu and Hamada[5] suggest plotting the residuals of all treatments

within each block and seeing if the pattern varies from block to block. They also point out that a "block-by-treatment interaction often suggests interesting information about the treatment and blocking variables", but they do not say why it is interesting.

Unfortunately, in the few practical applications that we found, the additivity assumption is not always checked. Moreover, not all reference texts emphasize the importance of checking it. It is as if it were impossible for blocks and treatments to interact, which is not the case. For example, Daniel[13] reanalyzes an experiment on growing mangolds that was conducted in four blocks of the same field and originally reported by Kempthorne[14], in which five variables are contemplated at two levels ($2^5$). An ingenious residual plot – physically placing the residual value in each block– shows a possible interaction between blocks and treatments undetected in the original analysis. As discussed in the next section, the existence of these interactions can be of extraordinary importance because they allow finding robust conditions.

Finally, we want to remind readers that blocked designs are a kind of split-plot designs. A block is essentially a set of runs conducted under more homogenous conditions than those outside the set, and this is also the definition of a whole-plot in split-plot designs. As pointed out by Box *et al.*[3] (p. 337): "The split-plot is like a randomized block design (with whole plots as blocks) in which the opportunity is taken to introduce additional factors". After saying that the models for analyzing split-plot and blocked designs are the same, Goos and Jones[15] (p. 240) state: "What is different between a split-plot experiment and a blocked experiment is the terminology used and the reason for using each type of experiment". In other words, the only difference between a blocked $2^{k-p}$ design and the corresponding split-plot design lies in the contrasts that are considered of interest.

# 3. Robustness

The aim of robust parameter design is to find the control factor settings that minimize system (product or process) variation due to changes in noise factors – factors that are hard to control and which cause variation in the response. These settings are sought through experimentation. Examples of noise factors frequently cited in the literature are environmental conditions and process variables that cannot be kept constant, such as differences in raw materials or in conditions under which the product will be used. Notice the similarity to blocking factors. This fact is made even clearer when looking at Table 2, which contains the examples used for presenting robust designs in the three books mentioned above.

*Table 2: Examples of robust designs in DOE industrial influential books*

| Book | Context | Noise factors | Reference |
|---|---|---|---|
| Box, Hunter and Hunter[3] | Medical packaging sealing | Environmental: sealing and dwell time | Not cited |
| | Industrial detergent | Customer process conditions: Temperature, water hardness and concentration | Based on Michaels[16] |
| Montgomery[4] | Leaf spring production | Process: quench oil temperature | Pignatiello and Ramberg[17] |
| | Development of an elastomeric connector | Conditioning environmental characteristics: time, temperature and humidity | Byrne and Taguchi[18] |
| Wu and Hamada[5] | Layer growth in wafer production | Location and facet | Kackar and Shoemaker[19] |
| | Leaf spring production | Process quench oil temperature | Pignatiello and Ramberg[17] |

In experiments aiming at finding robust conditions, there are two basic situations. In the first situation, the noise factors are unknown and we measure the output over time or location in order to "see" the variability provoked by the important noise factors. Asilahijani et al.[20] call these "robustness experiments". In the second situation, we know the noise factors whose effect we want to mitigate and, therefore, we can change them deliberately in order to induce variation in the response. These are called "desensitization experiments" in Asilahijani et al.[20]. It is well known that the second case is much more efficient than the first (Steinberg and Bursztyn[21]). In robustness experiments, we are forced to study the summary statistics of the different responses obtained for each experimental condition. Although there are other proposals like signal to noise ratios, we generally analyze two statistics: one to account for the level (typically the average) and another to control the variability (typically the standard deviation or a transformation of it). This approach has the added inconvenience that it does not provide information on the cause of the variability being neutralized, which makes it

difficult to detect cause-and-effect mechanisms. In desensitization experiments, we deal directly with the response, and robust conditions are found by studying the control-by-noise interactions. In fact, when noise factors affect the output, it is only possible to achieve robustness if these control-by-noise interactions exist. In this paper, given the similarity between blocking and noise factors, desensitization experiments are the ones of interest.

It is sometimes said that there are two ways to arrange control and noise factors in a design matrix: the so-called product array and the combined array. However, it is very easy to realize that it is always possible to write any design both ways. For convenience, we will now use the product array notation. It consists of setting up one design for the control factors and another one for the noise factors, then running combinations of both designs, generally all combinations. For economic reasons, restrictions are frequently imposed on the randomization of the design, thus making it a split-plot design (Box and Jones[22], Jones and Nachtsheim[23]). Depending on how the runs are randomized, Box, Hunter and Hunter[3] distinguish four different ways to run these designs. As discussed below, these methods have implications on the analysis and the ease of execution. The four methods are:

a)  Fully randomized. As stated before, this method is generally uneconomical and complicated to run.

b)  As a split-plot, with the control factors as whole-plots and the noise factors as subplots. On many occasions, this is the easiest way to run the experiment. It takes advantage of the split-plot arrangement for estimating the noise main effects as well as the noise-by-noise and noise-by-control interactions with the subplot variance. Unfortunately, this estimates the control main effects and control-by-control interactions with the whole-plot variance.

c)  As a split-plot, with the noise factors as whole-plots and the control factors as subplots. This should be the method of choice whenever possible. It takes full advantage of the split-plot arrangement and estimates control main effects, control-by-control interactions and noise-by-control interactions, all contrasts of interest, with subplot variance. Unfortunately, from the point of view of running the experiments it is sometimes unpractical.

d)  As a strip-block. This method has the advantage of a greatly reduced number of runs. However, it is inconvenient in that it implies three sources of error and it is very difficult to assess significance unless the number of control and noise factors is large.

Note that case c), the most highly recommended, is the same split-plot design that appears when blocking a design. There are only two changes: one is minor and only regards nomenclature; the other is conceptual and very important. The minor change is

6

that which is called a blocking factor in blocked designs is here called a noise factor, and that which is called simply a factor is here called a control factor. The conceptual change considers that it is possible to find robust conditions if control and noise factors interact, so these interactions are indeed of great interest.

# 4. Block and Robust Designs: A Unified Approach

To summarize the two preceding sections, we emphasize here that:

1.  The type and nature of noise and blocking factors is very similar and often equal.

2.  Both blocking and robust designs aim to neutralize variability. The difference is that the objective in blocking designs is limited to avoiding the influence of blocks in the estimation of factor effects; while the objective in robust designs is more ambitious and, in addition to offsetting the influence of noise factors (blocks) on the estimation of the effects, the aim is to neutralize the variability they provoke by studying their interactions with control factors.

3.  Blocked designs are split-plot designs

4.  There is a blatant contradiction between considering blocks to be additive when blocking and hoping that noise factors will interact with control factors when looking for robust conditions.

The first point, which is conceptual, has already been widely discussed in the previous sections, using for argument's sake the examples of Wu and Hamada[5]; Box, Hunter and Hunter[3]; and Montgomery[4]. A further indication that their nature is the same is provided by the fact that some authors call blocking factors "noise factors" (Anthony[24]) or "nuisance factors" (Montgomery[4]), although they still consider blocking and robustness to be separate topics.

Point two has to do with the objectives and, consequently, with the design types. The objective of robust design –to find robust conditions by studying whether there are important noise-by-control factor interactions– is complementary to the blocking objective and does not undermine its attainment in any way. We believe that, in practically all industrial situations, robustness is a desirable characteristic.

Point three addresses design types. We have seen in Section 2 that blocked designs are split-plots, and in Section 3 we have commented that it is often very convenient to run robust designs as split-plots, especially when the hard-to-change factors (whole plots) are the noise factors (situation c). Notice that, in both situations, the design is a split-plot with noise/blocking factors as whole plots.

Together with the objective discussion, the blatant contradiction in point four provides justification for unifying both approaches.

To conclude, we believe that the concept of blocking factor must be replaced by the concept of noise factor and, consequently, robustness should always be considered an objective. The only exception to this would be when blocking arises in the context of sequential experimentation. Naturally, the importance given to robustness will have to be balanced with the importance attributed to the control factor effects and the experimental design should reflect this balance.

In what follows, we discuss how to choose the levels of the noise factors and how to select experimental designs that would allow experimenters to balance good control factor estimation and finding robust conditions in accordance with their interests. Special attention is given to the case of having more than two blocks.

### *Choosing noise/blocking factor levels*

In robust experiments, the criterion for choosing noise factor levels is simple. They must be far enough apart to cause a noteworthy – even exaggerated – variation in the response. This exaggeration will facilitate the identification of noise-by-control interactions and, thus, will allow finding robust conditions.

When blocking a design, we are concerned with identifying the blocking factors, not their levels. Consider typical blocking factors such as those in Table 1: operator, shift, day, raw material batch and spindle. If we are concerned only with eliminating their effects from our estimation of control factors effects, their levels do not have to be chosen because they come naturally. However, if we are also trying to identify robust conditions, we can choose the levels appropriately and take advantage of the fact that, in split-plot designs, the size of the difference among whole plots does not in any way affect the estimation of the subplot effects or the subplot by whole plot interactions. For example, we can select the operators with more different outcomes or we can select raw material batches that are either very near or even slightly outside both extremes of the tolerance interval. A somewhat similar idea is presented by Daniel[13] under the name of *blocking for generality*, when he states on page 66: "Our results would be as precise, and our base broader, if we plan to take half the data on one batch and half on a batch separated as widely as possible from the first", or by Browne, MacKay and Steiner[24] in assessing measurement systems.

### *Selecting the design*

Bisgaard[25] discussed at length the issue of choosing robust designs by taking advantage of the fact that they are split-plot designs. He reminds us of the concept of split-plot confounding that was introduced by Bartlett[26] in 1935 and explains its importance in

robust designs. The idea is simple: split-plot confounding arises when the generators of a fractional two-level split-plot design mix whole plot and subplot factors. Aside from being a very interesting possibility for robustness, this is exactly what it is done when blocking $2^k$ designs. So, once again, we see that block and robust designs belong to the same family. Let us now consider a very simple example from the point of view of blocking and robustness.

Take, for example, a $2^3$ design in two blocks. The most reasonable as well as standard design uses $B_1 = ABC$ as the block generator because, assuming block additivity, the main effects and two-factor interactions are then free from confounding. (Figure 1). Notice that the generator mixes whole plot ($B_1$) and subplot ($ABC$) factors; thus, the design has split-plot confounding. This is always the case when blocking.



Figure 1: Design matrix and alias structure for a $2^3$ design in two blocks.

If, instead of thinking in two blocks, we think that we have one noise factor (W) and are interested in robustness, then the best option is to use $C = AB$ as a generator. In this way, we can estimate the noise-by-control interactions free of confounding. The trade-off is that the control main effects are confounded with the two-factor interactions. (Figure 2).



Figure 2: Design matrix and alias structure for an 8 run design with 3 control factors and 1 noise factor.

Which of the two designs is best? The answer depends on three considerations:

- The priorities and objectives of the experiment. If we are more interested in robustness than in the effects of control factors on the response level, the second design is preferable. However, if we prioritize finding the effects of control factors while keeping an eye on the noise-by-control interactions, the first alternative may be preferable.

- The experimenter's prior knowledge about the effect of factors and possible interactions.

- The possibilities of conducting further experiments using a sequential strategy.

Obviously, as the number of control and noise/block factors increases, the number of design options increases, which opens a wide range of possibilities for balancing different objectives.

***Designs divided into two blocks***

Designs divided into two blocks are completely equivalent to robust designs with one noise factor in which the noise factor is the hard-to-change factor.

If we consider eight-run designs divided into two blocks, the number of design alternatives is really small. Table 3 presents the most convenient generators for all possible eight-run designs and for both situations: when the priority is estimating the effects of control factors (the customary blocking situation) and when the priority is finding robustness (assuming the main interest is in control-by-noise factor interactions). It can be seen that: in two cases (five and six control factors), the generators are exactly the same; in one, they are the same but with different assignation (four control factors); and the two generators are different only when there are three control factors (which is the case discussed above).

*Table 3: Eight-run experiments divided into two blocks. Choice of generators depending on priority*

| Control factors | Number of Blocks/Noise factors* | Priority | |
| :---: | :---: | :---: | :---: |
| | | Control factor estimation | Robustness |
| 3 | 2 / 1 | W=ABC | C=AB |
| 4 | 2 / 1 | D=ABC | D=AB |
| | | W=AB | W=ABC |
| 5 | 2 / 1 | D=AB, E=AC, W=BC | |
| 6 | 2 / 1 | D=AB, E=AC, F=BC, W=ABC | |

*We denote W as the noise/blocking factor

As we have seen, the experimenter does not have many choices in eight-run designs. Basically, there exist only one or two possibilities for each combination of number of runs, number of control factors and number of noise or blocking factors. However, as the number of runs increases, the number of possibilities grows substantially. Even for 16-run designs, the experimenter can choose from among many designs that which best fits his or her particular problem.

For example, consider the case in which we want to study 5 control factors (A, B, C, D and E) and one noise/blocking factor (W) in 16 runs. Now, imagine that we prioritize the estimation of the control factor effects and are willing to assume that W is very unlikely to interact with control factors. Then, a good option is to use E=ABCD and W=AB as generators (Situation 1 in Table 4). If we are much more interested in finding robust conditions, a better choice of generators is then D=AB and E=AC, those of a $2^{5-2} \times 2^1$ product matrix (Situation 2 in Table 4), notice that it in this design all control-by-noise two-factor interactions are not confounded with main effects or other two factor interactions. As noted by Bisgaard[25], the use of split-plot confounding provides a wide range of intermediate situations. For example, in our case, using E=BCD and W=ABC as generators provides estimates of all main effects, design factors as well as noise factors, free of aliasing from any two factor interactions; while two-factor interactions involving a noise factor are confounded with other two-factor interactions (Situation 3 in Table 4).

*Table 4: Confounding patterns for three different sets of generators for a 16-run design with 5 control factors and one noise/blocking factor*

| Situation 1 | Situation 2 | Situation 3 |
|---|---|---|
| E=ABCD & W=AB | D=AB & E=AC | E=BCD & W=ABD |
| W+AB | W | W |
| A+BW | A+BD+CE | A |
| B+AW | B+AD | B |
| C | C+AE | C |
| D | D+AB | D |
| E | E+AC | E |
| AC | BC+DE | AB+DW |
| AD | BE+CD | AC+EW |
| AE | AW | AD+BW |
| BC | BW | AE+CW |
| BD | CW | BC+DE |
| BE | DW | BD+CE+AW |
| CD+EW | EW | BE+CD |
| CE+DW | | |
| DE+CW | 2 d.f. to estimate noise | 2 d.f. to estimate noise |

### Designs divided into more than two blocks

It is easy to see that, when a design is divided into two blocks, the blocking factor can be assimilated to a noise factor and that its interactions with control factors can be used to attain robust conditions. Things are a bit more complicated when the design is divided into 4 or more blocks. To show that this is also possible to assimilate blocking and noise factors when the design is divided in more than two blocks, consider a $2^4$ in four blocks and let us call the four factors A, B, C, and D. The standard block generators (see, for example, Box[3] or Wu[5]) are: $B_1=BC$, $B_2=ABD$ and $B_3=ACD$ (forget for the moment the aliasing caused by this choice). Table 5 shows the design matrix and the contrasts used as block generators.

*Table 5. Block generator contrasts for a $2^4$ design divided into four blocks*

| Blocks | A | B | C | D | $B_1$ (BC) | $B_2$ (ABD) | $B_3$(ACD) |
|--------|-----|-----|-----|-----|------|------|------|
| I   | 1  | 1  | -1 | -1 | -1 | -1 | 1  |
| I   | -1 | -1 | 1  | -1 | -1 | -1 | 1  |
| I   | -1 | 1  | -1 | 1  | -1 | -1 | 1  |
| I   | 1  | -1 | 1  | 1  | -1 | -1 | 1  |
| II  | -1 | -1 | -1 | -1 | 1  | -1 | -1 |
| II  | 1  | 1  | 1  | -1 | 1  | -1 | -1 |
| II  | 1  | -1 | -1 | 1  | 1  | -1 | -1 |
| II  | -1 | 1  | 1  | 1  | 1  | -1 | -1 |
| III | -1 | 1  | -1 | -1 | -1 | 1  | -1 |
| III | 1  | -1 | 1  | -1 | -1 | 1  | -1 |
| III | 1  | 1  | -1 | 1  | -1 | 1  | -1 |
| III | -1 | -1 | 1  | 1  | -1 | 1  | -1 |
| IV  | 1  | -1 | -1 | -1 | 1  | 1  | 1  |
| IV  | -1 | 1  | 1  | -1 | 1  | 1  | 1  |
| IV  | -1 | -1 | -1 | 1  | 1  | 1  | 1  |
| IV  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |

Looking at these contrasts it is easy to see that:

$B_1$ measures the difference between blocks I and III and blocks II and IV.
$B_2$ measures the difference between blocks I and II and blocks III and IV.
$B_3$ measures the difference between blocks I and IV and blocks II and III.

As a simple example, imagine now that the response for such an experiment is the one provided in Table 6. In this case, the only effects different from zero are the blocking factor $B_1$ and the $AB_1$ interaction; and both effects are equal to $b$. Figure 3 (a) shows the expected difference between blocks I and III and blocks II and IV; and Figure 3 (b) shows that, when A is at low level, the response is the same in the four blocks.

*Table 6. Control-by-block interaction compensating for differences between blocks in a four-block design*

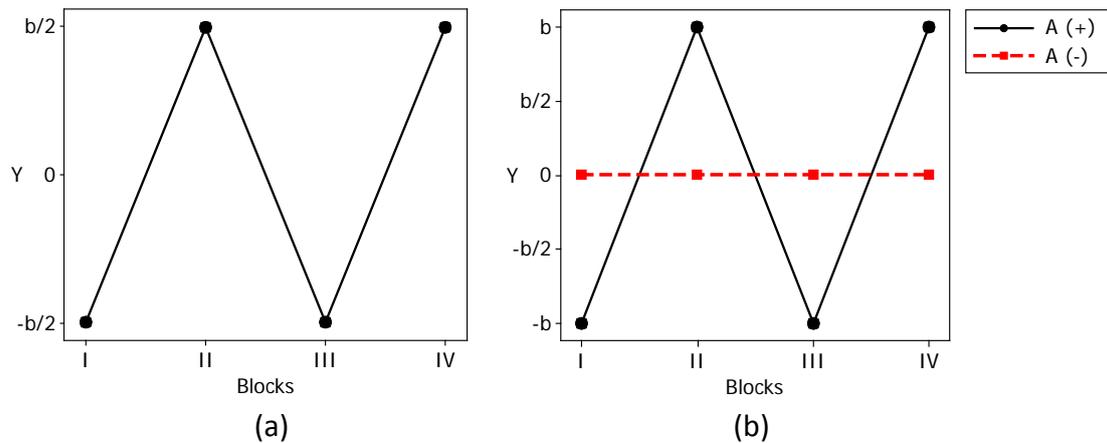| Blocks | A | B | C | D | $B_1$ (BC) | $AB_1$ (ABC) | Y |
|---|---|---|---|---|---|---|---|
| I | 1 | 1 | -1 | -1 | -1 | -1 | $-b$ |
| I | -1 | -1 | 1 | -1 | -1 | 1 | 0 |
| I | -1 | 1 | -1 | 1 | -1 | 1 | 0 |
| I | 1 | -1 | 1 | 1 | -1 | -1 | $-b$ |
| II | -1 | -1 | -1 | -1 | 1 | -1 | 0 |
| II | 1 | 1 | 1 | -1 | 1 | 1 | $b$ |
| II | 1 | -1 | -1 | 1 | 1 | 1 | $b$ |
| II | -1 | 1 | 1 | 1 | 1 | -1 | 0 |
| III | -1 | 1 | -1 | -1 | -1 | 1 | 0 |
| III | 1 | -1 | 1 | -1 | -1 | -1 | $-b$ |
| III | 1 | 1 | -1 | 1 | -1 | -1 | $-b$ |
| III | -1 | -1 | 1 | 1 | -1 | 1 | 0 |
| IV | 1 | -1 | -1 | -1 | 1 | 1 | $b$ |
| IV | -1 | 1 | 1 | -1 | 1 | -1 | 0 |
| IV | -1 | -1 | -1 | 1 | 1 | -1 | 0 |
| IV | 1 | 1 | 1 | 1 | 1 | 1 | $b$ |



*Figure 3: Block effect of $B_1$ (a) and A by block interaction (b)*

In fact, a suitable interaction between $B_1$ and any control factor A, B, C or D could compensate for differences between blocks I and III and blocks II and IV. Similarly adequate interactions between $B_2$ or $B_3$ and A, B, C or D would allow compensating for differences between the corresponding block pairs. Thus, interactions between $B_1$, $B_2$ or $B_3$ and A, B, C or D could be useful for increasing robustness and, hopefully, a combination of them could make the process completely insensitive to variation between blocks.

Take, for example, the case reported by Young[12] on the "Effort to close the hatch in a new car design" and included by Wu and Hamada (Table 1) as an exercise. The experiment studies the impact of three hatch production factors: Length (L), Position (P) and Firmness (F) on the effort to close the hatch. The design is a replicated $2^3$ design in four blocks of four runs. Cars were the blocking factor, because it was known that there was car to car variation. The results in Young's experiment show that the only significant effect is the difference between blocks (Figure 4).

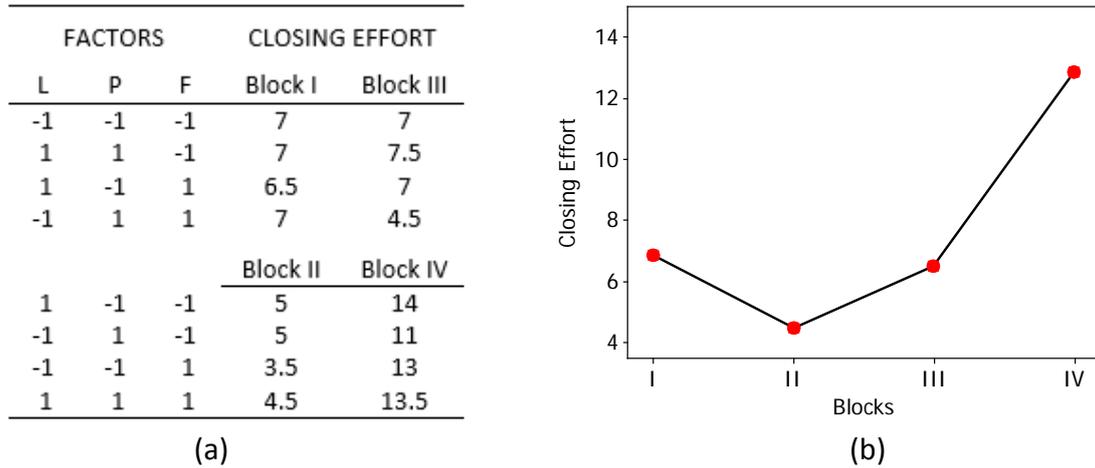| FACTORS | | | CLOSING EFFORT | |
|---|---|---|---|---|
| L | P | F | Block I | Block III |
| -1 | -1 | -1 | 7 | 7 |
| 1 | 1 | -1 | 7 | 7.5 |
| 1 | -1 | 1 | 6.5 | 7 |
| -1 | 1 | 1 | 7 | 4.5 |
| | | | Block II | Block IV |
| 1 | -1 | -1 | 5 | 14 |
| -1 | 1 | -1 | 5 | 11 |
| -1 | -1 | 1 | 3.5 | 13 |
| 1 | 1 | 1 | 4.5 | 13.5 |

(a)



(b)

*Figure 4: Design and response of Young Experiment (a) and block effects (b)*

Unfortunately, in this case there are no interactions between cars (the noise/blocking factor) and control factors. However, if the results were to have been those shown in Table 8, the differences between blocks would have been the same; but, thanks to the interactions between F and the three blocking factors, they are offset by working with F at low level (Figure 5).
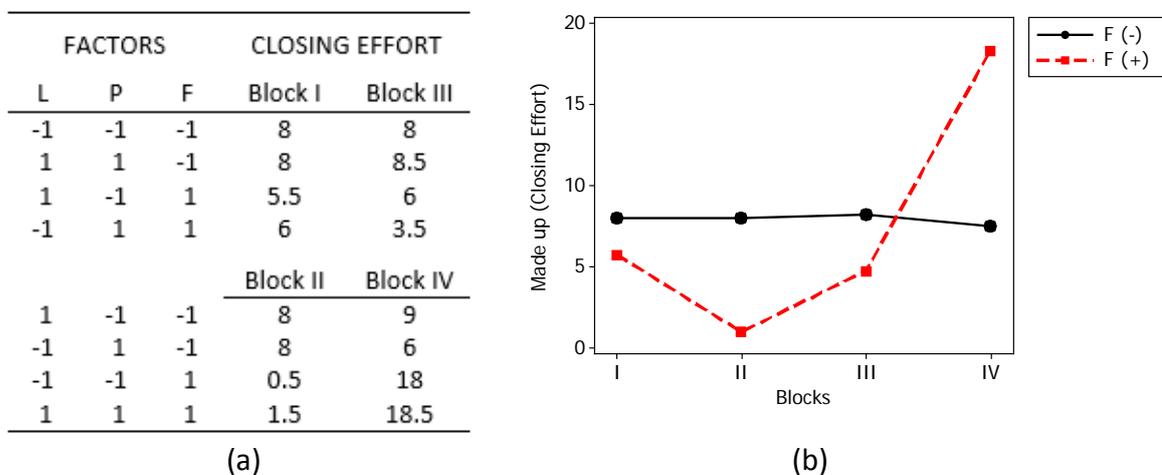
| FACTORS | | | CLOSING EFFORT | |
|---|---|---|---|---|
| L | P | F | Block I | Block III |
| -1 | -1 | -1 | 8 | 8 |
| 1 | 1 | -1 | 8 | 8.5 |
| 1 | -1 | 1 | 5.5 | 6 |
| -1 | 1 | 1 | 6 | 3.5 |
| | | | Block II | Block IV |
| 1 | -1 | -1 | 8 | 9 |
| -1 | 1 | -1 | 8 | 6 |
| -1 | -1 | 1 | 0.5 | 18 |
| 1 | 1 | 1 | 1.5 | 18.5 |

(a)



(b)

*Figure 5. Design and made up response of Young Experiment (a) and F by block interaction in this case (b).*

Just imagine how happy the engineers would have been if they would have found robust production conditions such as these, which would have made the effort of closing the hatch independent of the car! Obviously, it would be fantastic to understand not only the technical reasons behind these interactions, but also what type of car variability (i.e., what car "noise factor" other than non-specified differences among cars) they are fighting. However, this is not necessary for realizing that putting Firmness (F) at a low level will reduce the Closing Effort variability.

The question is: If any of these interactions had existed, would the engineers have noticed them? Probably not. To detect them, two things are needed: a design that allows their estimation and an appropriate analysis. In this case, the design is a split-plot in which cars (blocks) are the whole plots and the control factor the subplots. The block generators used were: $B_1$=replicates and $B_2$=LPF. Table 7 shows that the alias structure allows the estimation of control-by-block/noise interactions. It also shows that if the experimenter is only interested in blocking, his or her vision of the alias structure is very different and it will be difficult to detect control-by-block/noise interactions, if they exist.

*Table 7. Young's design alias structure, depending on the approach, robustness or blocking*

| Robustness Non-additive blocks | Traditional Blocking Additive blocks |
|---|---|
| I | I |
| L | L |
| P | P |
| F | F |
| $B_1$ | LP |
| $B_2$ | LF |
| $B_3$ | PF |
| $LB_1$ | 6 error d.f. |
| $PB_1$ | |
| $FC_1$ | |
| $LB_2 + PF$ | |
| $PB_2 + LF$ | |
| $FB_2 + LP$ | |
| $LB_3$ | |
| $PB_3$ | |
| $FB_3$ | |

In the $2^4$ four-block design at the beginning of this section, we deliberately avoided talking about the alias structure in order to center our attention on the issue of robustness. For robustness purposes, we have used the standard block generators that provide a poor alias structure (see Table 8). Of course, this means that they are also very

poor for checking block additivity, an assumption that, once again, cannot be taken for granted.

*Table 8. Alias structure for the $2^4$ in four-block design. Standard generators*

| | |
|---|---|
| I | |
| A | |
| B + CB$_1$ | |
| C + BB$_1$ | |
| D | *In the example and discussion above,* |
| B$_1$ + BC | *we have deliberately used the AB$_1$* |
| B$_2$ | *interaction because is free from* |
| B$_3$ | *aliasing.* |
| AB + DB$_2$ | |
| AC + DB$_3$ | *Unfortunately, any other choice of* |
| AD + BB$_2$ + CB$_3$ | *generators provides a similar or worse* |
| BD+ AB$_2$ | *alias structure.* |
| CD+ AB$_3$ | |
| BB$_3$ + CB$_2$ | |
| AB$_1$ | |
| DB$_1$ | |

As we have seen in the two blocks case, the experimenter can generally choose from among several options the design generators that best fit his or her interest, specifically in terms of control factor estimation or robustness. However, in the three-factor eight–run and in the four or more factors and sixteen-run cases, the number of choices is really limited. Furthermore, in these scenarios the aliasing is so strong that – no matter which generators are chosen– it is almost impossible to estimate control factor by block interactions or to check the block additivity assumption.

Naturally, when the relationship between the number of control factors and the number of runs improves, the number of design options increases and there appear many situations in which the experimenter can prioritize control factor estimation or robustness. For example, this happens when there are three control factors and 16 runs (as in Young's case) or when the experimenter can afford 32 runs to study 4 or 5 factors in 4 blocks.

## 5. Conclusions

When conducting experiments in industrial environments, it is almost always a better option to consider as noise factors rather than blocking factors the factors that cannot

16

This allows pursuing two objectives: obtaining homogeneous experimental conditions (good estimation of the effects of control factors); and, if we are lucky and noise factors interact with control factors, making the system insensitive to nuisance factors under "normal" conditions.

From a theoretical standpoint, there is no difference between considering as blocking or noise factors the non-controllable factors that introduce undesired variability. In both cases, the design is a split-plot design and it is analyzed accordingly. One advantage of this unified approach is that the use of split-plot confounding allows the designer a wider option of designs, depending on his or her objectives. The experimenter is not confronted with two Manichean choices: putting all the experimental effort in estimating control factor effects (assuming additivity between blocking/noise and control factors) or putting all the effort in estimating control by noise interactions (in detriment of the estimation of control factor effects). The wider choice of designs is perfect for obtaining the maximum benefit from taking into account technical knowledge and also from using sequential experimentation.

Finally, we believe that if this unified approach is presented in design of experiments courses that are aimed at industrial applications in both professional and university settings, it will facilitate the understanding of the topic and consequently will help the industrial community in designing better experiments. In our view, the only reason for presenting the two topics separately is historical.

# 6. Acknowledgments

# References

1.  Box GEP. Statistics for Discovery. *Journal of Applied Statistics*2001;**28**:285-299.

2.  Vining, G. Technical Advice: Design of Experiments, Response Surface Methodology, and Sequential Experimentation. *Quality Engineering*2011;**23**:217-220.

3.  Box GEP, Hunter JS, Hunter WG. Statistics for Experimenters: Design, innovation and discovery. Wiley:Hoboken, 2005.

4. Montgomery DC. Design and Analysis of Experiments. Wiley:Hoboken, 2013.

5. Wu CFJ, Hamada M. Experiments: Planning, Analysis, and Parameter Design Optimization. Wiley: New York, 2009.

6. Bisgaard S. Industrial use of statistically designed experiments: Case study references and some historical anecdotes. *Quality Engineering*1992;**4**:547-562

7. Prvan T, Street DJ. An annotated bibliography of application papers using certain classes of fractional factorial and related designs. *Journal of Statistical Planning and Inference*2002;**106**:245–269.

8. Ilzarbe L,Alvarez MJ,Viles E, Tanco M.Practical applications of Design of Experiments in the field of engineering. A bibliographical review. *Quality and Reliability Engineering International*2008;**24**:417-428.

9. Sharif KM, Rahman MM, Azmir J, Mohamed A, Jahurul MHA, Sahena F, Zaidul ISM. Experimental design of supercritical fluid extraction – A review. *Journal of Food Engineering* 2014; **124**:105-116

10. Sanders D, Leitnaker MG, McLean RA. Randomized Complete Block Designs in Industrial Studies. *Quality Engineering*2001;**14**:1-8

11. Jensen WA, Kowalski SM. Response Surfaces, Blocking and Split Plots: An Industrial Experiment Case Study. *Quality Engineering*2012;**24**:531–542.

12. Young JC. Blocking, replication and randomization—The Key to effective experiments: A case study. *Quality Engineering*1996;**9**:269-277.

13. Daniel C. Applications of Statistics to Industrial Experimentation. Wiley:New York, 1976

14. Kempthorne O. Design and Analysis of Experiments. Wiley:New York, 1952

15. Goss, P. and Jones, B. (2011). Optimal Design of Experiments. A Case Study Approach. Chichester (UK): Wiley

16. Michaels SE. The Usefulness of Experimental Designs. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*1964;**13**:221-235.

17. Pignatiello JJ, Ramberg JS. Discussion of "Off-line quality control, parameter design, and the Taguchi method" By Kackar, R.N., *Journal of Quality Technology*1985;**17**:198-206

18. Byrne DM, Taguchi G.The Taguchi Approach to Parameter Design. *Quality Progress*1987; December, pp. 19-26

19. Kackar RN, Shoemaker AC. Robust Design: A Cost-Effective Method for Improving Manufacturing Processes. *Bell Labs Technical Journal*1986;**65**:39-50.

20. Asilahijani H, Steiner SH, MacKayRJ. Reducing Variation in an Existing Process with Robust Parameter Design. *Quality Engineering*2010;**22**:30-45.

21. Steinberg DM, Bursztyn D. Dispersion Effects in Robust-Design Experiments with Noise Factors. *Journal of Quality Technology*1994;**26**:12-20.

22. Box GEP, Jones S. Designing products that are robust to the environment. *Total Quality Management*1992;**3**:265-282.

23. Jones B, Nachtsheim CJ. Split-Plot Designs: What, Why, and How. *Journal of Quality Technology* 2009; **41**:340-361

24. Browne R, MacKay J, Steiner S.  Leveraged Gauge R&R Studies. *Technometrics* 2010; **52**: 294-302

25. Bisgaard S. The Design and Analysis of $2^{k-p} \times 2^{q-r}$ Split Plot Experiments. *Journal of Quality Technology*2000;**32**:39-56

26. Bartlett MS. Discussion of Complex Experiments by F. Yates. Supplement to the *Journal of the Royal Statistical Society II*1935; **2**:224-226

**Brief biography of each author**

**Pere Grima** is an Associate Professor at Universitat Politècnica de Catalunya-BarcelonaTech (UPC), where he also obtained his PhD. His research interests are quality management and design of experiments. He has more than twenty years of experience in helping companies implement statistical methods for quality control and improvement. He has been an advisor for the European Quality Award and has acted as a consultant to several multinational companies in Six Sigma projects.

**Lluís Marco-Almagro** is an Associate Professor at Universitat Politècnica de Catalunya-BarcelonaTech (UPC), where he teaches and conducts research on topics related to industrial statistics and quality improvement. He holds a degree in industrial engineering and a PhD in statistics. One of his research areas of interest is design of experiments and statistical process control. He has also worked on statistical methods for emotional industrial design. He has served as an expert in helping many companies with the design and analysis of industrial experiments. He has also trained Black Belts and served as a consultant in Six Sigma projects over the last 10 years.

**Xavier Tort-Martorell** is an Associate Professor at Universitat Politècnica de Catalunya-BarcelonaTech (UPC). His research interests and industrial consulting activities cover quality management and improvement, Six Sigma, design of experiments and data-based decision making. He is currently an Associate Editor for TQM&BE (Total Quality Management & Business Excellence), QTQM (Quality Technology and Quantitative Management), and for the International Journal of Quality and Service Sciences. He has

been an advisor for the European Quality Award, a member of the jury for the Catalan Quality Prize and was president of the European Network for Business and Industrial Statistics (ENBIS) in 2012 and 2013.