

Design of Experiments (DOE)—A Valuable Multi-Purpose Methodology

Miryam Barad

Department of Industrial Engineering, Faculty of Engineering, Tel Aviv University,
Tel Aviv, Israel
Email: barad@post.tau.ac.il

Received 3 May 2014; revised 6 June 2014; accepted 18 June 2014

Copyright © 2014 by author and Scientific Research Publishing Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY).
<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

The DOE methodology is an effective tool for upgrading the level of measurement and assessment. In any design, planning or control problem the designer is faced with many alternatives. He/she is challenged to develop design approaches that can meet both quality and cost criteria. The way experiments are designed greatly affects the effective use of the experimental resources and the easiness with which the measured results can be analyzed. This paper does not present new evidence based on designed experiments. Its objective is solely to show how useful application of multifactor experiments is in a variety of circumstances and decision making scenarios. The paper reviews three published examples where this method was used in different contexts: quality control, flexible manufacturing systems (FMS) and logistics systems. The physical experiment has been carried out to improve the quality of a special type of batteries. The simulation experiment has been carried out to investigate the impact of several flexibility factors in a flexible manufacturing system. The numerical value of a complex analytical expression representing a customer oriented logistics performance measure has been calculated for different values of its parameters, *i.e.* the given numerical values of the investigated factors. It enabled a methodical examination of all factor effects and especially their interactions, thus shedding light on complex aspects of the logistics decision problem. In these examples, cases from different contexts were presented, enabling to view design of experiments as a powerful ingredient for improving decision making in a variety of circumstances.

Keywords

Design of Experiments, Confounding Designs, Deterministic Response, Quality Control, Flexible Manufacturing Systems, Logistics Systems

1. Introduction

Design of experiments is a multi-purpose technique [1]. Its usage is not limited to physical experiments but can be applied to simulation experiments [2], to the investigation of calculated results of complex analytical expressions whose parameters are methodically varied or to other decision problems, where the effects of several factors are examined. The origins of DOE were in the UK, where at the beginning of the previous century the method was developed by Sir Fisher and was successfully used in agriculture. However, its application in western industries was meager. “Taguchi methods” that were developed in Japan during the 1960’s and were disseminated in the west during the 1980’s represent a specific style of experimental design with specific advantages and pitfalls [3] [4].

Applying DOE to physical or to simulation experiments as well as to other purposes, consists of some well-defined activities such as, selecting relevant measures of the system effectiveness (the response, or dependent variable), selecting the factors (independent variables) to be varied and the levels to be investigated [5].

DOE is an *economical technique*, because reliable results can be obtained based on a relatively small number of observations. Correctly analyzing data from a small but well-designed experiment will yield statistically sound interpretations. Its great advantage is a methodical examination of the factor effects and especially of their interactions that may shed light on more complex aspects of a decision problem. Interaction between factor effects is a basic, important concept in the experimental design methodology. Interaction between two factors occurs when the effect of one factor, as estimated at a given level of another factor, is not the same at all levels of the other factor and vice-versa.

It seems of interest to mention that *interactions* (positive or negative) between factor effects, is sometimes suggested by system thinkers, for whom the quantitative aspect of the concept might be less clear, as a factor able to provide great organizational benefits [6].

This paper does not present new evidence based on designed experiments or any standard statistical analysis technique such as analysis of variance. Its objective is solely to show how useful application of multifactor experiments is in a variety of circumstances and decision making scenarios.

The paper reviews three examples. They are set in industrial or logistics environments. They were published some years ago in different contexts: quality control, flexible manufacturing systems (FMS) and logistics systems. Each example is related to a different application category: a physical experiment, a simulation experiment and a complex analytical expression. The design of the physical and the simulation experiment employed fractional factorial design, a compact but confounding design, which results in aliases. The physical experiment has been carried out to improve the quality of a special type of batteries. It had resolution IV, where it is not possible to avoid confounding two factor interactions. The original paper emphasized that in such compact designs reliable previous information is necessary in order to avoid erroneous conclusions. The second example deals with a simulation experiment. Simulation is a computer experimentation technique that uses input/output models for gaining information about a real-world system or process. Hence, as part of any simulation project, experiments have to be designed so as to yield the desired information needed for meeting the project’s objectives. The scene of the simulation experiment was set in a flexible manufacturing environment and its aim was to investigate the impact of several factors in this environment. It had resolution V, where no two factors interactions are confounded with each other and thus all two factor interactions could be investigated. The third example examines the numerical value of a complex analytical expression representing a customer oriented logistics performance measure, as calculated for different values of its parameters (the given numerical values of the investigated factors). Its design was a full factorial with 5 factors and enabled a methodical examination of all factor effects and especially their interactions, thus shedding light on some complex aspects of the logistics decision problem.

The paper is organized as follows. Section 2 describes the physical experiment whose goal was to improve the performance of a nickel cadmium battery, published in *Quality Progress* [7]. This example is the most detailed one here because of the special circumstances related to the analysis of its results. It advised against relying on outcomes based on very compact fractional designs, when not enough reliable information on basic hypotheses is known. Section 3 summarizes the analysis of the simulation experiment whose goal was to investigate the impact of some flexibility factors in a flexible manufacturing system. It was published in *International Journal of Production Research* [8]. Here, the objective of this example is to emphasize the capability of multi-factor experiments to investigate the effects of factor interactions for improving the design and operation of (flexible) manufacturing systems. Section 4 examines the effect of several factors on a customer oriented, calculated, logis-

tics performance measure. It was published in the International Journal of Production Economics [9]. This example shows how experimental design can be used when the response, or dependent variable, is not a stochastic variable, but represents a calculated numerical value of an expression which due to its complexity cannot be examined analytically. The paper ends with a discussion presented in Section 5.

2. A Fractional Factorial Design of a Physical Experiment

Fractional factorial design is very economical but it should be used carefully. Research must take place ahead of time to determine which variables can be ignored. Preparation prevents incorrect assumptions that lead to erroneous conclusions. In the company where the case study took place the experimental design technique was to be tried through physical experiments, intended to improve the performance of a special type of nickel-cadmium battery. The problem that arose regarded some changes to be made in battery design and manufacturing to improve capacity at low temperature. Thus the effects of several factors on the battery capacity had to be evaluated.

2.1. Selection of the Factors to Be Investigated and Their Levels

The product development engineers agreed to investigate six factors, with each factor being run at two levels. A full factorial experiment would require 64 observations. Because of the budgetary and time constraints it was decided to start with a condensed pilot experiment involving one-quarter of the full experiment, *i.e.* 16 observations. Such designs are discussed in several textbooks (see e.g. [5])

The factors selected for investigation were:

Factor **A**—type of plate. The levels of factor A were: A1, negative plate and A2, positive plate.

Factor **B**—additions to electrolyte. The two levels of factor B were: B1, no addition to the electrolyte, B2, a given mixture and concentration of additions.

Factor **C**—manufacturing method. Two methods for manufacturing plates, method I, C1 and method II, C2 were to be compared to determine their effect on capacity.

Factor **D**—impurities in the electrolyte. The two levels of factor D were: D1, no impurities, D2 a given impurities concentration.

Factor **E**—testing temperature. This factor was of particular interest in the experiment because of the intended usage conditions of the battery. The low level of this factor, E1 was -20°C and the high level, E2 was 20°C .

Factor **F**—discharge rate. F1 and F2 represented respectively, the low and high rate of discharge and were respectively selected as half and twice the number of ampere hours of the battery's nominal capacity.

The response, *dependent variable*, was the capacity expressed in ampere hours per plate unit area.

2.2. Choice of the Design

When designing a fractional factorial experiment, not all the interactions can be evaluated. In an experiment designed as one-quarter of a complete factorial, groups of four different interactions are actually considered as one because their individual effects cannot be separated. If any of them has a significant influence on the measured dependent variable (here, the capacity), the statistical analysis will eventually show it but will be unable to point to the source. They will be confounded. The confounded effects are called aliases.

The fractional factorial design in this study was based on an a priori requirement to analyze and estimate the main factor effects of each of the six factors and also each of three two factor interactions (BC, CD and EF), considered by the product development engineers as the most likely to affect the battery capacity. They considered all the other two factor interactions to be insignificant. Another objective of the experiment was to supply an estimate of the variance caused by randomness.

The testing conditions expressing the selected design appear in **Table 1**.

Each row represents a treatment combination whose capacity had to be measured in the experiment, thus providing one observation.

Sixteen standard area plates were prepared: eight negatives and eight positives. Four of the negatives and four of the positives were manufactured by method I. Four of the negatives and four of the positives were manufactured by method II. Each group of four plates of the same type (negative or positive) manufactured by the same method (I or II) comprised a physical experimental unit. One of the four was randomly selected to be immersed

Table 1. The fractional factorial design represented by the 16 detailed combinations.

Type of Plate	Manufacturing	Electrolyte	Temperature	Discharge Rate
Negative	Method I	KOH	-20°C	Low
Negative	Method I	KOH + AD	+20°C	Low
Negative	Method I	KOH + IMP	-20°C	High
Negative	Method I	KOH + AD + IMP	+20°C	High
Negative	Method II	KOH	+20°C	High
Negative	Method II	KOH + AD	-20°C	High
Negative	Method II	KOH + IMP	+20°C	Low
Negative	Method II	KOH + AD + IMP	-20°C	Low
Positive	Method I	KOH + IMP	+20°C	Low
Positive	Method I	KOH + AD + IMP	-20°C	Low
Positive	Method I	KOH	+20°C	High
Positive	Method I	KOH + AD	-20°C	High
Positive	Method II	KOH + IMP	-20°C	High
Positive	Method II	KOH + AD + IMP	+20°C	High
Positive	Method II	KOH	-20°C	Low
Positive	Method II	KOH + AD	+20°C	Low

in a standard volume of electrolyte with no additions or impurities, another was immersed in electrolyte with additions, a third was immersed in electrolyte with impurities and the remaining one was immersed in electrolyte with additions and impurities. The particular temperature and discharge rate for each immersed plate were according to the table. The testing order of the 16 observations was random as well. As mentioned earlier, because of the fractional structure of the experiment confounded effects or aliases were created. The main issue in the design of a fractional experiment is to enable the separation, through statistical analysis, of the effects that were considered important prior to designing the experiment. That means not including two or more such important effects in the same group of aliases. If a certain alias group is found to have a significant effect, it is attributed solely to the main effect or the particular interaction previously considered important within the group. The aliases emerging from the designed experiment here are listed in [Table 2](#). This table does not include the 4 defining contrasts, namely, $I = ABCE = ACDF = BDEF$, which define the design resolution and cannot be analyzed (see 2.4 below and [5] for explanations). The six main effects and the three factor interactions considered important prior to the designing of the experiment are underlined. There is not more than one underlined set of letters in the same row. Some of the rows do not contain any underlined sets, meaning that no important effect is comprised therein. Based on the design assumptions, the effects of these rows will serve to estimate the variation caused by randomness. The significance of the important effects selected to be investigated will be separated as a result of the statistical analysis and tested against the estimate of the random variation.

2.3. Analysis of the First Data Set

The findings, based on the statistical data analysis, showed that the factors that significantly affected the battery capacity were: **F**—discharge rate, **E**—temperature and **BC**—interaction between the addition to the electrolyte and the manufacturing method. The evaluated standard deviation caused by randomness was 0.51 ampere-hour/plate unit area.

The product design engineers were reluctant to accept all the results of the experiment. They expected some imbalance between the positive and the negative plates, *i.e.* that the effect of factor **A** would be significant. But, as mentioned above, the data analysis did not find the effect of factor **A** significant. Also, it was hard to logically

Table 2. The list of the confounded effects or aliases.

<u>A</u>	BCE	CDF	ABDEF
<u>B</u>	ACE	ABCDF	DEF
<u>C</u>	ABE	ADF	BCDEF
<u>D</u>	ABCDE	ACF	BEF
<u>E</u>	ABC	ACDEF	BDF
<u>F</u>	ABCEF	ACD	BDE
AE	<u>BC</u>	CDEF	ABDF
AF	BCEF	<u>CD</u>	ABDE
BD	ACDE	ABCF	<u>EF</u>
AB	CE	BCDF	ADEF
AC	BE	DF	ABCDEF
AD	BCDE	CF	ABEF
BF	ACEF	ABCD	DE
ABD	CDE	BCF	AEF
ABF	CEF	BCD	ADE

explain the highly significant effect of BC, the interaction between the addition to the electrolyte and the manufacturing method.

But, luckily, at this stage we got supplementary data on the already conducted experiment. It should be remembered that the experiment comprised 16 physical experimental units that had to be tested at a certain temperature and discharge rate, as detailed in [Table 1](#). In practice, each designed capacity test was preceded by a required preliminary capacity test, conducted for all units at normal conditions, high temperature level (E2) and low discharge rate (F1). After that each physical unit was recharged and retested as designed. These preliminary 16 observations were not designed but were performed because of technical requirements. They represented a full factorial with 4 factors A, B, C and D (the first 3 columns of [Table 1](#)), executed at constant values of the remaining factors E and F (high temperature and low discharge rate).

2.4. Analysis of the Second Data Set and Some Concluding Remarks

The analysis of the second data set revealed that factor **A** (type of plate) was the most significant. Factor **C** (manufacturing method) was also significant and so was their interaction, **AC**. At the low rate of discharge (F1) and high temperature (E2) the estimated standard deviation caused by randomness was 0.11 ampere hour/plate unit area, much lower than its previous estimate (0.51).

Integrating the findings of the two data sets, several conclusions could be reached

- The main effects of factors A and C, which were significant at normal testing conditions (low rate of discharge and high temperature level) were obscured and thus insignificant when the results were averaged over the entire range of test conditions imposed on the originally designed experiment. The results of the second data set indicated that the negative plates provided significantly more capacity under normal test conditions (as the product engineers expected). But they were affected more than the positives under severe test conditions.
- The statistical analysis of the first data set found the interaction BC very significant. By examining its aliases ([Table 2](#)) it was seen that the aliases group comprising interaction BC also comprised interaction AE, between the type of plate and temperature. The meaning of this significant interaction is that the temperature effect is not the same for the two types of plates.
- Looking at the combined findings of the two data sets it can be seen that this result fits as a piece in a puzzle. In designing the experiment AE was not considered an important interaction. In its alias group the only con-

sidered interaction was BC. Therefore, the highly significant effect that was attributed to BC (and was hard to explain) was actually the effect of AE. They both reduced the evaluated capacity.

- The apparent inconsistency of results with regard to C (manufacturing method) between the two data sets could be explained similarly. Presumably the interactions CE and CF, initially assumed to be negligible, were actually important. They could not be tested because they were grouped with the estimation of the error variance. Some confirmation of this hypothesis is provided by the significant increase of the error standard deviation when carrying out the experiment at varying discharge rate and temperature (0.51) as compared to its estimate under normal testing conditions.
- Fractional factorial design is an efficient technique. It can analyze the effects of many factors using a low number of observations. Here, 16 observations were used to investigate the effects of 6 factors. However, it should be used with caution. Reliable prior information theoretical and/or empirical is necessary to avoid erroneous conclusions as illustrated by the analysis here of the first data set.
- Defining contrasts represent the core of the fractional factorial design. They define the design resolution. The length of the defining contrast determines its resolution (see [5]). Here the defining contrast group was: $I = ABCE = ACDF = BDEF$. Accordingly, its resolution was IV and all the six main effects were free, *i.e.* not confounded with other main effects or with two factor interactions. However, the latter were confounded with other two factor interactions.

3. A Fractional Factorial Design of a Simulation Experiment

This second example also discusses a fractional factorial design, but its resolution was V. In this resolution not only the main factor effects are free but the two factor interactions as well. In other words they are not confounded with other main effects or with two factor interactions but solely with higher order interactions. The design of the simulation experiment is detailed in **Table 3**. Also detailed here (see Section 3.2) is the usage of batches, a specific feature of a simulation experiment which improves its efficiency. However, most of the experiment results are not presented, because they are related to FMS research and not relevant here. The presented results are related to significant interaction effects and are meant to emphasize the capability of a designed experiment to investigate interaction effects.

Table 3. The fractional factorial design (failures not considered).

Machine versatility	Mixes	Control methods	Utilization level
Low (R1)	Mix 1 (E1)	none	High
		B2	Low
		C2	Low
		B2 and C2	High
	Mix 2 (E2)	none	Low
		B2	High
		C2	High
		B2 and C2	Low
High (R2)	Mix 1 (E1)	none	Low
		B2	High
		C2	High
		B2 and C2	Low
	Mix 2 (E2)	none	High
		B2	Low
		C2	Low
		B2 and C2	High

3.1. The Factors to Be Investigated and the Design of the Simulation Experiment

One of the original paper objectives was to examine the impact of six factors on each of two performance measures of a simulated FMS. Five factors were examined at two levels each and the sixth factor was examined at three levels.

The factors were as follows.

R—a design factor representing *system* versatility (R1 low level, R2 high level)

B, C—two control factors, expressing scheduling rules assigning priorities to the parts waiting to be processed based on the available information. Factor **B** is part oriented, meant to prevent long delays, by assigning a higher priority to a part according to its waiting time in system. Factor **C** is system oriented, meant to avoid machine starvation. It assigns a higher priority to a part, if its next processing machine is idle. Each rule may be considered or not considered, a factor being at its respective high level ($B2 = 1, C2 = 1$) when the rule is considered, otherwise ($B1 = 0, C1 = 0$).

U—an environmental factor, planned machine utilization ($U1 = 0.60, U2 = 0.80$).

E, F—two factors representing changes. Factor **E** expresses alterations in type of demand and its two levels are two different mixes generated from an aggregate mix: E1 a mix with skewed load distribution, E2 a mix with uniform load distribution of the various part types.

The sixth factor, **F**, stands for failures, which were considered at three levels: no failures, 5% failures, 10% failures. The percentages express the ratio MTTR/MTBF (Mean Time to Repair/Mean Time between Failures).

A complete five factor experiment each at two levels involves 32 treatment combinations. We used a fractional factorial experimental design ($1/2 2^5$), resulting in 16 combination. The design of the 16 combinations used a defining contrast of 5 letters, $I = REBCU$. Accordingly, its resolution was **V**, enabling to free all main effects and two factor interactions. Each of the 16 combinations was tested at three failure levels.

Two on-line performance measures were considered:

1) Td_i —Standardized tardiness, a unified service-oriented measure defined as “total waiting time in system of part type $i, i = 1, 2, \dots, n$, expressed as a ratio of its processing time”;

2) Work in process (WIP), a design oriented measure.

The system was investigated under unlimited storage and the values of the planned machine utilization were selected in such a way as to ensure that each machine is below saturation ($U_{max} < 1$). Under these conditions, work in process has been selected as a performance measure able to supply relevant information at the design stage due to its expected sensitivity to failures and to the planned utilization level.

3.2. Simulation Analysis

The simulation model and the experimental conditions were developed using Siman network modeling. A simulation run represented a particular treatment combination. Length of each simulation run was 14,400 minutes (240 hours).

The analysis was focused on the steady-state result, considering the principles prevailing in a “non-terminating” simulation environment. The Siman output processor was extensively utilized. Through the “filter” command, the transient period for each run was determined, the transient results were deleted and the remaining steady state period was divided into batches. The means of the batches approximately represent independent observations (see [2] for a more detailed description). Statistical testing of the independence between batches (Fisher’s method) is performed automatically by the output processor.

Lengths of the transient periods for runs with no equipment failures, varied between 10 and 15 hours. Under the same conditions, the size of the independent batches varied between 80 and 140 consequent observations (part units manufactured).

The equipment failure conditions were modeled by assuming an exponential distribution for both the time between failures and the repair duration. A mean time between failure, $MTBF = 600$ minutes was assumed, coupled with a mean time to repair, $MTTR1 = 30$ minutes and $MTTR2 = 60$ minutes for 5% and 10% failures respectively. Lengths of the transient periods under equipment failures exceeded 40 hours.

It is worth mentioning that if instead of batches replicated runs had been used, a minimum of 120 hours would have been required for one replication. When 9 independent replications per run (equivalent to our nine batches) are considered, the total simulated time per run becomes: $9 \times 120 = 1080$ hours as compared to our 240 hours per run. Hence a relative efficiency of more than 4/1 was obtained.

It is beyond the scope of the article to present the standard technical statistical analysis of the results. It seems though appropriate to mention that the means of the batches were used in the analysis of the multi factor experiment for estimating the factorial effects. Variances between the batches means within any run were pooled, leading to an overall estimate of the white noise (random effect) variance. The statistical significance of the effects (F-test) was performed relative to the above variance of the random effect.

3.3. Some Results

The following partial results of the above simulation experiments are presented as examples of the designed experiments capability to investigate interaction effects (positive or negative) between factors.

- Under no failures, the most important counter effect (leading to a decrease in standardized tardiness and WIP) was the interaction **RU** (negative). Its interpretation is that a high system versatility level (**R2**), is especially effective under high planned machine utilization (**U2**). The second best counter effect was the main effect of factor **R** (system versatility) meaning that it improved the system performance under all operating conditions. The interaction **EU** (negative) was also found significant. It implies that the system can better cope with a more uniform mix (**E2**) and that this is especially important at a high utilization level (**U2**).
- When 10% failures were considered, the most important effect in reducing standardized tardiness and WIP was the interaction **BU** (negative). This implies that under equipment failure, the on-line control strategy assigning scheduling priority by tardiness (factor **B** at its high level **B2**) is especially effective when the system is operated at a high planned utilization level (**U2**).

4. A Full Factorial Design to Investigate a Calculated Complex Analytical Expression

The third example is definitely different. It does not describe a physical or a simulation experiment but uses design of experiments to investigate the effects of several factors and their interactions on a calculated complex expression, representing the dependent variable. Accordingly, the dependent variable (or response) is not a stochastic variable but a deterministic one. It represents a customer oriented logistic performance measure, denoted in the original paper, logistics dependability. The scenario of this example was framed in an inventory system. Logistics dependability emphasized quick response delivery through reduced supply uncertainty during an inventory replenishment system. It was defined as the probability that there is no supply shortage until the end of the replenishment cycle, for given starting condition and inventory policy

4.1. The Design and Its Rationale

One aim of the original paper was to get a better understanding of the logistic decision problems and their eventual solutions. To achieve that aim it was necessary to examine the effects of several factors on the logistics dependability of a given logistic system. As the complexity of the equations for calculating logistics dependability (see original paper) did not permit an analytical approach, the paper showed how the statistical experimental technique could be used in an innovative manner to easily obtain effective results. Several factor types that may affect dependability were considered. *Design* factors which could be manipulated in order to increase the system dependability; *environmental* factors which were selected in order to represent different normal operating conditions. As the paper dealt with flexibility, a factor representing *changes* was also selected. Such factor is particularly relevant for emphasizing the value of flexibility, represented in the paper by trans-routing flexibility, a design factor. It was measured as u —the number of transshipment links per end user at the echelon level. In a rigid system which has no transshipment links, $u = 0$. In a flexible system with N end users at the same echelon level, maximal trans-routing flexibility is obtained for $u = N - 1$, meaning that all end users at the same echelon level are inter-linked.

The experiment was designed as a full two level factorial experimental design in five factors, 2^5 . The investigated system consisted in 4 end users, $N = 4$.

The two *design* factors were:

A—Trans-routing flexibility supporting decision making in real time, measured by u ($u = 0$ —low level, $u = 3$ —high level);

B—Service level supporting static planning based decisions, measured by $1 - \alpha$ (0.8—low level, 0.9—high

level).

The two *environmental* factors were:

C—Lead time, measured in terms of a fraction of the replenishment cycle time L ($L/8$ —low level, $L/4$ —high level);

D—Demand variability, measured by variability coefficient σ/μ (0.25—low level, 0.30—high level);

E—a *change* factor representing increase in the demand variability during the replenishment cycle (no change—low level, a given change—high level).

The following questions were investigated in the original paper:

- How effective is trans-routing flexibility, as compared to an increased service level, for coping with changes that occur during a replenishment cycle?
- Is trans-routing equally effective for any demand variability?
- Does lead time affect the benefits of trans-routing?

These questions were answered through an analysis of the interaction effects between the investigated factors.

4.2. Analysis and Results

Typically, in a statistical experimental design the response is a stochastic variable. Eventually, the statistical analysis of the experimental results is used to estimate variances and to build confidence intervals for some unknown parameters. As mentioned before, in this example the response (dependability) was not a stochastic variable, but was calculated for varying values of the parameters as detailed above. The role of experimental design was to discriminate in a methodical way between the important factor effects and the negligible ones. As the analysis results are not used to estimate variances, or to build confidence intervals, we deemed it was legitimate to use this technique.

The data analysis in the original paper combined ANOVA (analysis of variance) with a graphical procedure relying on a Normal Probability plot. The procedure is described in the experimental design literature (see e.g. [5]). According to this procedure standardized effects are calculated for each factor and each interaction. The Normal Probability Plot is used to graphically estimate the important effects. The effects that are negligible are normally distributed with mean zero and white noise variance and will tend to fall along a straight line, whereas significant effects will have non-zero means and will not lie along the straight line. The white noise variance needed for the ANOVA is estimated by pooling the variances of the negligible effects that lie along the straight line. Here the white noise variance stands for the negligible factor effects and no further usage will be made of its numerical value. The results in terms of analysis of variance (ANOVA) were as follows.

The zero p-values in ANOVA disclosed the important factor effects. These were: the two *design* factors, namely **A**—trans-routing flexibility and **B**—service level, as well as **E**—the change in the demand variability. Several interaction effects, namely **AD**, **AE** and **BE**, were also important. Among the factors exhibiting a zero p-value we deliberately ignore the effect of factor **D**—demand variability. In this particular case the mean effect of factor **D**, which proved to be “significant” was actually the effect of factor **D** for a flexible system. In other words, the effect of factor **D** was only present when factor **A** was at a high level and thus the analysis only considered the effect of interaction **AD**. When there is no flexibility, *i.e.* factor **A** is at a low level, the service level neutralizes the effect of any increase in the demand variability (see original paper for further explanations and calculations).

The zero p-values do not show the effect direction, *i.e.* whether it has a positive effect, *i.e.* it increases dependability or it has a negative effect on dependability. The Normal Probability Plot supplies this information.

Four factor effects were positively placed with respect to the line and two were negatively placed. The two most significant positive effects were the effects of **A** and **B**, the *design* factors. Let us recall that the effect of factor **A** represented the increase in dependability obtained for a system with highest flexibility, $u = N - 1 = 3$, as compared to a rigid system, $u = 0$. The effect of factor **B** represented the increase in dependability following a decrease in risk, from 0.2 to 0.1 (an increase in the service level $1 - \alpha$ from 0.8 to 0.9). As judged by their magnitude these two effects were equivalent. The two other positive effects were the two interactions, **AD** and **AE**. The most negative factor effect was the effect of factor **E**, the change in the demand variability that increased the risk level by 0.043. The next negative effect was the interaction **BE**. The lead time did not affect the logistic dependability.

Let us now interpret the interaction results.

The negative interaction between the change factor **E** and the design factor **B**, means that under changing conditions increasing the service level is not an effective procedure. By contrast, trans-routing flexibility is more effective under changing conditions. This is illustrated by the positive effect of interaction **AE**.

Also, as shown by the positive significant effect of interaction **AD**, trans-routing flexibility (factor **A**) is more effective for higher demand variability than it is for lower demand variability. This result can be explained by the fact that high demand variability necessitates a high stock level. This stock is not well utilized in a rigid system. By contrast, a flexible system utilizes this higher stock in a dynamic way during a replenishment cycle thus improving the system dependability.

5. Discussion

The reviewed examples show that Design of Experiments is a multi-purpose technique that can be applied in a variety of circumstances where a methodical examination of factor effects and especially their interaction effects are vital in order to shed light on complex aspects of a decision problem.

The multi-factor physical experiment showed that prior research is critical in concise fractional factorial experiments. The simulation experiment was able to provide answers to all the research questions asked in relation to design oriented decisions in a flexible manufacturing system. Designing, performing and analyzing the simulation runs as a multi-factor experiment made them efficient. DOE enabled to analyze the effects of several factors and their interaction on the calculated complex expression of the defined logistics system.

The supplied information on factor effects and on factor interactions effects provided an insight view and a better understanding of all the examined practical decision problems.

References

- [1] Box, G.E., Hunter, W.G. and Hunter, J.S. (2005) *Statistics for Experimenters*. 2nd Edition, Wiley, New York.
- [2] Law, A.M. and Kelton, W.D. (2000) *Simulation Modeling and Analysis*. 3rd Edition, Mc Graw Hill, New York.
- [3] Taguchi, G. (1986) *Introduction to Quality Engineering*. Asian Productivity Organization, Tokyo.
- [4] Dehnad, D. (ed.) (1989) *Quality Control, Robust Design and the Taguchi Method*. Wadsworth & Brooks/Cole, Pacific Grove, California.
- [5] Montgomery, D.C. (1997) *Design and Analysis of Experiments*. 4th Edition, Wiley, New York.
- [6] Dean, J.W. and Bowen, D.E. (1994) Management Theory and Total Quality: Improving Research and Practice through Theory Development. *Academy of Management Review*, **19**, 392-418.
- [7] Barad, M., Bezalel, C. and Goldstein, J.R. (1989) Prior Research Is the Key to Fractional Factorial Design. *Quality Progress*, **22**, 71-75.
- [8] Barad, M. (1992) The Impact of Some Flexibility Factors in FMSs—A Performance Evaluation Approach. *International Journal of Production Research*, **30**, 2587-2602. <http://dx.doi.org/10.1080/00207549208948178>
- [9] Barad, M. and Even-Sapir, D. (2003) Flexibility in Logistics Systems—Modeling and Performance Evaluation. *International Journal of Production Economics*, **85**, 155-170. [http://dx.doi.org/10.1016/S0925-5273\(03\)00107-5](http://dx.doi.org/10.1016/S0925-5273(03)00107-5)

Scientific Research Publishing (SCIRP) is one of the largest Open Access journal publishers. It is currently publishing more than 200 open access, online, peer-reviewed journals covering a wide range of academic disciplines. SCIRP serves the worldwide academic communities and contributes to the progress and application of science with its publication.

Other selected journals from SCIRP are listed as below. Submit your manuscript to us via either submit@scirp.org or [Online Submission Portal](#).

