Some Thoughts About Your DOE Project

(30% Weighting)

TOR: On Stream
Due: 22 May 2015 (5pm)
Type: Individual Project
Report Style: Lab Report (structure given)

Task/Aim

To design and analyse a multifactor experiment of the form 2^k (WK6) with blocks or 2^{k-p} (WK7) and make actionable conclusions.

Two factor experiments are ***not*** allowed; they are too simplistic. Also, you can limit the number of trials to 16 due to the time factor in fairly tedious experiments.

Some Options:

- ✓Three factor full factorial experiments (**2 blocks** min).
- ✓Four factor full factorial experiments (**2 blocks** min).
- ✓Five (or more) factor *fractional factorial experiments with adequate resolution (Res IV or **V**).* √

A fractional factorial with blocking (in general) gives a very complex alias structure. Hence should be avoided.

Consider simple homemade experiments that you like E.g.

Title	Factors	Response
Studying the accuracy of throwing darts	Hand (L/R) Distance to the Board Lighting	Accuracy (the mean distance from the bull's eye via repeated measurements)
Studying the time to boil water **Not a good project at all!**	Pan size (Small/Large) Lid (Off/On) Quantity (1 lit/1.5 lit) Salt (without/with)	Time to boil (seconds)
Studying the height rise of scones	Butter (low/high) Flour (low/high) Milk (low/high)	The rise (mm)
Studying the flight time of a paper aeroplane	Wing Length (S/L) Wing Width (S/L) Wing Shape (1/2) Paper Type (80/100gsm) Load (No/1 paper clip)	Time in flight (s) **Do not plagiarize annotated examples!**

1. Introduction (10%)

State the problem (preferably in the form of a research question) and the associated process succinctly (the problem should not be too trivial!).

E.g. of a reasonably nontrivial research question

> "To what extent does the distance, the level of lighting, and the hand used affect throwing accuracy in a game of darts, for a normal (non-ambidextrous) person?"

Identify the experimental factors and the response variable/s. *Note: No need to study two response variables (e.g. Y and the STD or range of Y) unless your research question or objectives warrants so.*

Blocking (if you are doing a full factorial experiment)

- Decide what your two blocks should be (since blocks represent a nuisance factor, they should not be part of your research questions/objectives).

- Do not go for more than 2 blocks.

- Be aware that there are 2 methods of blocking—confounding a higher order interaction with blocks, and block on replicates—and follow the most logical: block on replicates (replicate 1 in block 1 and replicate 2 in block 2). You would like a kill 2 birds with one stone!

- When you tell Minitab that you want 2 replicates and 2 blocks, it will always give a block-on-replicate design! Why ☺?

Try to avoid hard to change factors. For example, among the four factors A (temp: 160^0C-180^0C), B (fan: on/off), C (time: 40-60 min), D (ingredient: low/high) used to study baking quality (e.g. softness), <u>factor A (temp.) may be a hard to change (time consuming) factor </u>(from low to high and vice versa)!

If you do not randomise runs as often as is necessary, strictly speaking, that has to be accounted for in your ANOVA (such designs, known as split-plot designs, not covered ☺).

StdOrder	RunOrder	A	B	C	D	Y (Softness 1-10)
15	1	-1	1	1	1	
10	2	1	-1	-1	1	
8	3	1	1	1	-1	
3	4	-1	1	-1	-1	
7	5	-1	1	1	-1	
2	6	1	-1	-1	-1	
5	7	-1	-1	1	-1	
9	8	-1	-1	-1	1	
14	9	1	-1	1	1	
11	10	-1	1	-1	1	
13	11	-1	-1	1	1	
4	12	1	1	-1	-1	
12	13	1	1	-1	1	
1	14	-1	-1	-1	-1	
16	15	1	1	1	1	
6	16	1	-1	1	-1	

2. Aim and objectives (10%)

One or typically a combination of these are typical…

a) To characterise the system (e.g. by writing down a regression equation/s to predict the response of interest) and determine the optimum settings for each factor, given its current range. **Mandatory!**

b) To identify practically significant main effects and interactions for the purpose of optimisation.

c) To reduce the variability of a certain process (e.g. kicking goals), whilst achieving another goal (maximise the number of kicks) . Note: 2 response variables (*Y* and the std of *Y*) required!

Aim and objectives contd./-

Your design resolution (e.g. Full, Res V, Res IV) should be adequate to fulfil your experimental objectives.

State the research hypotheses and/or expected results.

H_0 and H_A are not research hypotheses although typically, H_A coincides with a research hypothesis (e.g. the hand used has an effect on throwing accuracy).

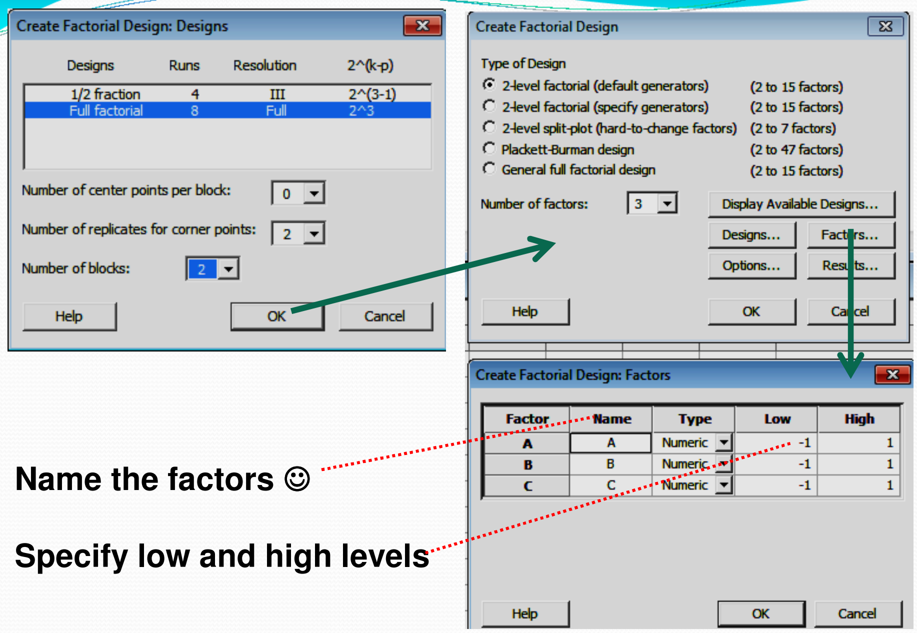
3. Materials and Method (25%)

- List the material used and describe how the experiment was conducted.
- Attach a photograph or two to show the experimental setup (colour printouts are not essential).

Creating Your Design Matrix

- Before you run the experiment ask Minitab to set up your design matrix (unless you tell otherwise, Minitab will give the runs in the random order, which is of course what you want). Then save your Project (Surname.MPJ) to conduct the experiment, to record response data, and to analyse the data.
- You need to upload your Minitab project file also!
- Consider naming the factors and specifying the low and high levels.

Useful Screen Dumps



Name the factors ☺

Specify low and high levels

Background variables

Background variables aka controllable nuisance factors: Very important to identify what these are. Describe what you did with the BVs. Did you:

(a) Hold these constant throughout the experiment? (then you can run the experiment in one block) or,

(b) Go for blocking such as ½ of the runs in Block 1 and the other ½ of the runs in Block 2. **Required anyway if you going for a full factorial!**

Note: (a) and (b) above require two different patterns of randomisation.

Nuisance Variables (aka uncontrollable nuisance factors)

There are literally infinite number of nuisance variables (Deming called these "common causes") that disturb your response variable/s (i.e. noise). Attempt to identify say, the three most potentially influential uncontrollable nuisance factors (e.g. depending on your experiment say wind, ambient temperature, humidity, measurement precision etc.).

Response variable/s (must not be binary!!)

If your response variable/s is a well defined measure (e.g. time, distance), the question of **reliability and validity** of Y does not arise—our predecessors have validated these measurement scales! E.g. Everyone knows what 50^0C is!

On the other hand, if you are measuring a new or unfamiliar or abstract variable such as the "taste", "baked quality" of a baked product (e.g. muffins), make sure that you define (operationalise) the concept appropriately for the purpose of measurement; e.g. you may say "6 people were asked to taste the product and rate it in a scale of 1 (horrible) to 10 (fantastic).

Type of design used should be fully described!

This should include, among other things (e.g. full factorial or fractional factorial, if fractional factorial the design generators, defining relation, the resolution, and the alias structure) whether you randomised all your runs completely (in which case you will have run the whole experiment in one block) or you had to randomise your runs within each block.

- Materials and Method is also the section in which you should list your table of factors and their two levels (this should be done before showing the design matrix). Make sure that you select a reasonably wide difference between your -1 level and the +1 level for each factor as otherwise, your analysis may fail to detect an effect (i.e. small effect sizes could become undetectable).

- Also mention how you tried to ensure the accuracy and precision of your measurements.

Randomisation, Replication and Blocking

This assignment tests your theoretical knowledge on DOE, your ability to use the theoretical knowledge in a practical situation, and your ability to present your study.

Randomisation, replication and blocking are the three key principles on which DOE is built.

So we expect you to demonstrate that you understand these principles. This is not to say that you always need to run your experiment in more than one block nor that you need to replicate a say 2^4 or a 2^{5-1} design. If you did not use blocking (e.g. for a fractional factorial) or replication, justify why you didn't use the said strategies.

4. Results and Discussion (40%)

- Showing Results – Similar to what you did in the lab with Minitab.
- Discussion/Analysis - Similar to what we did in the lab class/theory class/tutorial class (mostly lab class!).
- Do not leave your results unexplained (e.g. unexplained cube plot, ANOVA table etc.).
- Final model, its adequacy based on the runs conducted (R^2, residual plots etc.).
- A discussion on how you could further confirm (validate) the adequacy of your model via confirmation runs. Do conformation runs if you can.

What if none of your factors were statistically significant?

This is OK as long as you did a good job of:

(a) stating plausible hypotheses;

(b) describing the experiment;

(c) making sure that your results won't get affected by significant systematic error and random error;

(d) analysing the data correctly;

(e) providing a strong conclusion section.

5. Conclusion (15%)

You need to state which of your hypotheses were supported by data (i.e. through your data analysis) and which weren't. You also need to show your final model. If applicable, you can also mention whether or not you were surprised with the findings.

Scope for further research (consider your study as part of a sequential learning process on the phenomenon you studied) and the limitations of your model should also be mentioned.