

An Online system  
for teaching the  
design (and analysis) of experiments

Carl James Schwarz  
Statistics and Actuarial Science  
Simon Fraser University  
Burnaby, BC, Canada V5A 1S6

[cschwarz@stat.sfu.ca](mailto:cschwarz@stat.sfu.ca)

All web pages are available at:

<http://www.stat.sfu.ca/~cschwarz/ExpDesign>

## Problem

Hunter (1977):

“In most courses on experimental design, students get no practice in *designing* experiments, although from home work assignments, they do get practice in *analyzing* data.”

Continuum of assignments/projects:

*Analysis: Home work in back of textbooks*

- data already collected
- all students get same correct answer
  
- some (unfortunate) learning outcomes:
  - \* one, unique answer to an experiment
  - \* estimates and standard errors don't vary

*Design: Actual Experiments*

- students design own experiment
- students collect own data
  
- some (unfortunate) outcomes:
  - \* costly
  - \* (ironically) too much variation
  - \* difficult to try different designs on same exp

## Solution

Hunter (1977):

Use simulated data where “each student receives data based upon the design he or she chooses ... if the teacher uses simulated data creatively, students can learn many important aspects of design and analysis in a most effective way... the untapped potential is enormous.”

Precedents (from survey sampling)

- StatVillage (Schwarz, 1997)
- SURVEY (Chang et al, 1992)

Desirable features for such a system:

- web based to make platform independent
- every student should get different responses
- multiple types of designs should be possible
- instructor can easily create and modify

# Solution: Student Perspective

## Initial Experimental Layout

### Is there an effect of masktype on time to complete a maze?

Do aromas affect the time to complete a task? In this experiment, you will investigate the effects of four different aromas (plain – the control; an apple scent, a banana scent; and a skunk scent) upon the time to complete a paper maze. You have pool of subjects (use their first names to identify them). Each test takes approximately 30 minutes so only eight tests can be done in any particular day.

#### How to run the experiment

Below you will find a field for the *Slot* and a pop-down item to set the factor level. Then press the *Run the experiment* button at the bottom of the page to have the experiment run and values returned to you. You may wish to print out your experimental design before submitting it to be run

Student Number:

Run	Day	Slot	Subject	Mask
1	1	1	<input type="text"/>	<input type="text"/>
2	1	2	<input type="text"/>	<input type="text"/>
3	1	3	<input type="text"/>	<input type="text"/>
4	1	4	<input type="text"/>	<input type="text"/>
5	1	5	<input type="text"/>	<input type="text"/>
6	1	6	<input type="text"/>	<input type="text"/>
7	1	7	<input type="text"/>	<input type="text"/>
8	1	8	<input type="text"/>	<input type="text"/>
9	2	1	<input type="text"/>	<input type="text"/>
10	2	2	<input type="text"/>	<input type="text"/>
11	2	3	<input type="text"/>	<input type="text"/>
12	2	4	<input type="text"/>	<input type="text"/>
13	2	5	<input type="text"/>	<input type="text"/>
14	2	6	<input type="text"/>	<input type="text"/>
15	2	7	<input type="text"/>	<input type="text"/>
16	2	8	<input type="text"/>	<input type="text"/>

Did you PRINT your design???

A file will be returned to you with the responses from the experiment.  
[Problems? Send a message to YourInstructor \(cschwarz@stat.sfu.ca\)](mailto:cschwarz@stat.sfu.ca)

Figure 1. The virtual laboratory prior to running an experiment. Students must enter their student number, complete the text boxes to identify subjects, and select the treatments to be allocated to subjects using the pop-down menu.

# Solution: Student Perspective

## Designing Experiment

### Is there an effect of masktype on time to complete a maze?

Do aromas affect the time to complete a task? In this experiment, you will investigate the effects of four different aromas (plain – the control; an apple scent, a banana scent; and a skunk scent) upon the time to complete a paper maze. You have pool of subjects (use their first names to identify them). Each test takes approximately 30 minutes so only eight tests can be done in any particular day.

#### How to run the experiment

Below you will find a field for the *Slot* and a pop-down item to set the factor level. Then press the *Run the experiment* button at the bottom of the page to have the experiment run and values returned to you. You may wish to print out your experimental design before submitting it to be run

Student Number:

Run	Day	Slot	Subject	Mask
1	1	1	carl	plain
2	1	2	lois	apple
3	1	3	matthew	skunk
4	1	4	marianne	banana
5	1	5	david	banana
6	1	6	dov	plain
7	1	7	connie	skunk
8	1	8	julie	apple
9	2	1	miriam	skunk
10	2	2	tim	plain
11	2	3	kim	plain
12	2	4	Fionna	banana
13	2	5	phillipa	apple
14	2	6	charmaine	banana
15	2	7	farell	skunk
16	2	8	sol	apple

Did you PRINT your design???

A file will be returned to you with the responses from the experiment.  
[Problems? Send a message to YourInstructor \(cschwarz@stat.sfu.ca\)](#)

Figure 2. A possible experimental design. This is a single factor completely randomized design (CRD). The student then clicks on the "Run" button to get the experiment data shown in Figure 3.

# Solution: Student Perspective Results from Experiment

Here are the results of your experiment.

Save this file as a text file using the File menu and the Save options.  
Then strip off the lines before and after the raw data.  
Be sure that the file is saved as a TEXT file.

```
Student number, Day , Slot , Subject , Mask , Response (Y)
999999999 01 01 carl plain 201
999999999 01 02 lois apple 183
999999999 01 03 matthew skunk 198
999999999 01 04 marianne banana 155
999999999 01 05 david banana 184
999999999 01 06 dov plain 163
999999999 01 07 connie skunk 173
999999999 01 08 julie apple 161
999999999 02 01 miriam skunk .
999999999 02 02 tim plain 107
999999999 02 03 kim plain 129
999999999 02 04 fionna banana 127
999999999 02 05 phillipa apple 140
999999999 02 06 charmain banana 125
999999999 02 07 farell skunk 110
999999999 02 08 sol apple 157
End of results listing
```

Figure 3. The results from the experiment run in Figure 2. If the same experimental design were to be run, different values would be returned. This file should be saved to a text file for further analysis by the student.

Solution: Student Perspective  
Possible Designs that can be run  
on this simple experiment

- able to run
  - CRD
  - RCB
  - GRCB
  - sub-sampling (pseudo-replication)
  - repeated measure

# Solution: Instructor Perspective Control File

Line	Control Field
1	Is there an effect of masktype on time to complete a maze?
2	Do aromas affect the time to complete a task? In this experiment, you will
3	investigate the effects of four different aromas (plain - the control; an apple
4	scent, a banana scent; and a skunk scent) upon the time to complete a paper maze.
5	You have pool of subjects (use their first names to identify them). Each test takes
6	approximately 30 minutes so only eight tests can be done in any particular day.
7	END of instructions - important that first 3 letters say END in uppercase.
8	1 - number of factors
9	4 Mask - number of levels for factor and name of the factor
10	plain 140 0 Then for each level, enter its label, the average,
11	apple 145 0 and the increase in the standard deviation
12	banana 140 0
13	skunk 155 2
14	2 8 - number of blocks and block size
15	Day - name of "blocks"
16	15 Std deviation of block effect
17	Slot - name of units within blocks
18	5 Std deviation of responses at unit level (MSE)
19	Subject - name of experimental unit
20	20 Experimental unit variation (subject effects)
21	.10 - fraction of missing values
22	InstructorName - Instructor name
23	cschwarz@stat.sfu.ca - instructor email address
24	no - trace flag (used for debugging) yes or no
25	yes - email flag (send student design to instructor via email?)

Figure 4. Control file used to generate the virtual laboratory in Figure 1.

Line 1: Title of the experiment  
 Lines 2-7: Explanatory text that appears on web page with designed experiment  
 Line 8: Number of factors  
 Line 9: The number of levels in the single factor and the factor name.  
 Lines 10-13: For each level, the level name, the level mean, and the increase in variation for that level.  
 Line 14: The number of blocks and block size.  
 Line 15: The name of the blocking variable.  
 Line 16: The standard deviation of the blocking effect.  
 Line 17: The name of the runs.  
 Line 18: The standard deviation of the run effect.  
 Line 19: The name of the experimental unit  
 Line 20: The experimental unit standard deviation.  
 Line 21: The fraction of missing values  
 Line 22: Instructor name and email address.  
 Line 23: The instructor's email address  
 Line 24: A trace flag used for debugging purposes – ordinarily not used.  
 Line 25: Does the instructor want the student design and responses emailed to them?

## Solution: Instructor Perspective

### Control file

- number of factors and names of factors
- number of levels/names of levels of factors
- number of blocks
- experimental units per block
  
- mean for each treatment combination
  
- block standard deviation
- run standard deviation
- experimental unit standard deviation
- excess treatment standard deviation
  
- MCAR rate

## Solution: Instructor Perspective: Statistical Theory

### Underlying statistical model

$$Y_{trt,block,subject,run} = \mu_{trt} + \mu_{block} + \mu_{subject} + \mu_{run} + \mu_{trt}$$

where

$\mu_{trt}$  - treatment mean

$\mu_{block}$  - block effect  $N(0, \sigma_{blocks}^2)$

$\mu_{subject}$  - subject effect  $N(0, \sigma_{subjects}^2)$

$\mu_{run}$  - run effect  $N(0, \sigma_{run}^2)$

$\mu_{trt}$  - excess variation for specific  
treatments  $N(0, \sigma_{trt}^2)$

Judicious choice of values leads to many different design characteristics.

# Solution: Student Perspective

## More Complex Design

### Is there an effect of masktype on time to complete a maze – II?

Do aromas affect the time to complete a task? In this experiment, you will investigate the effects of four different aromas (plain – the control; an apple scent, a banana scent; and a skunk scent) and gender (male or female) upon the time to complete a paper maze. You have pool of subjects (use their first names to identify them). Each test takes approximately 30 minutes so only eight tests can be done in any particular day.

#### How to run the experiment

Below you will find a field for the *Slot* and a pop-down item to set the factor level. Then press the *Run the experiment* button at the bottom of the page to have the experiment run and values returned to you. You may wish to print out your experimental design before submitting it to be run

Student Number:

Run	Day	Slot	Subject	Gender	Mask
1	1	1	carl	male	plain
2	1	2	lois	female	apple
3	1	3	matthew	male	plain
4	1	4	marianne	female	skunk
5	1	5	david	male	skunk
6	1	6	samantha	female	apple
7	1	7	meghan	female	skunk
8	1	8	paul	male	apple
9	2	1	joan	female	banana
10	2	2	mark	male	banana
11	2	3	shawna	female	plain
12	2	4	stephen	male	skunk
13	2	5	carolyn	female	plain
14	2	6	lynn	female	banana
15	2	7	richard	male	plain
16	2	8	stephen	male	banana

Did you PRINT your design???

A file will be returned to you with the responses from the experiment.  
[Problems? Send a message to YourInstructor \(cschwarz@stat.sfu.ca\)](#)

Figure 5. A possible experimental design for a two factor experiment. This is a two factor completely randomized design (CRD). The student clicks on the "Run" button to get the experimental data.

## Solution: Student Perspective More Complex Design

- able to run
  - CRD
  - RCB
  - GRCB
  - sub-sampling (pseudo-replication)
  - repeated measure
  - split-plot design

### References:

Chang, T. C., Lohr, S. L., and McLaren, C. G. (1992). Teaching Survey Sampling Using Simulation. *American Statistician*, 46, 232-237.

Hunter, W. G. (1977). Some Ideas About Teaching Design of Experiments, with 2<sup>5</sup> Examples of Experiments Conducted by Students. *American Statistician*, 31, 12-17.

Schwarz, C. J. (1997). StatVillage: An On-Line, WWW-Accessible, Hypothetical City Based on Real Data for Use in an Introductory Class in Survey Sampling. *Journal of Statistical Education [Online]* 5(2).