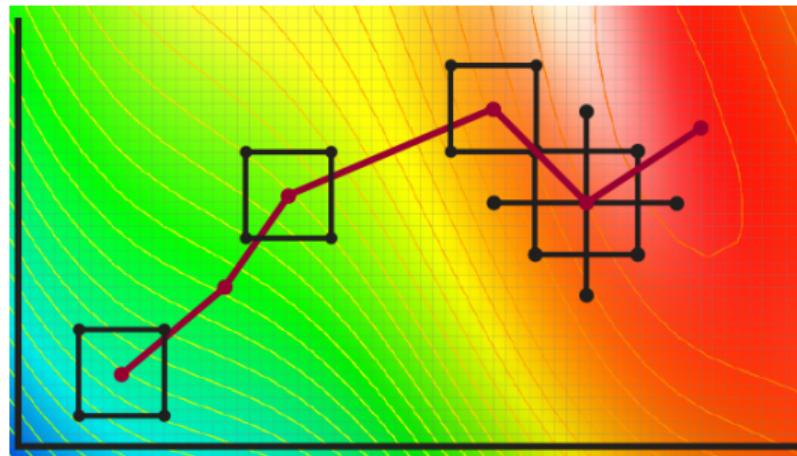


Experimentation for Improvement



© Kevin Dunn, 2015
<http://learnche.org/>

Design and Analysis of Experiments

Copyright, sharing, and attribution notice

This work is licensed under the Creative Commons Attribution-ShareAlike 4.0 Unported License. To view a copy of this license, please visit <http://creativecommons.org/licenses/by-sa/4.0/>



This license allows you:

- ▶ **to share** - to copy, distribute and transmit the work, including print it
- ▶ **to adapt** - but you must distribute the new result under the same or similar license to this one
- ▶ **commercialize** - you are allowed to use this work for commercial purposes
- ▶ **attribution** - but you must attribute the work as follows:
 - ▶ "Portions of this work are the copyright of Kevin Dunn", or
 - ▶ "This work is the copyright of Kevin Dunn"

(when used without modification)

The variables affecting our experimental system

We may classify them in several ways:

- ▶ those we know about, and those that are unknown
- ▶ variables we can control, and uncontrolled variables: *we'll consider these today*
- ▶ some we can measure, and others we cannot measure: *we'll consider these today*

Controllable and measurable variables and the need for randomization

Use a familiar example: ginger biscuits

We have 3 factors:

1. **A**: baking temperature
2. **B**: amount of baking soda
3. **C**: baking time



Order of experiments*:

Experiment	A	B	C	Time finished
5	—	—	+	08:32
1	—	—	—	09:46
8	+	+	+	10:50
3	—	+	—	12:05
6	+	—	+	13:16
2	+	—	—	14:30
4	+	+	—	15:57
7	—	+	+	17:09

* note the randomized order

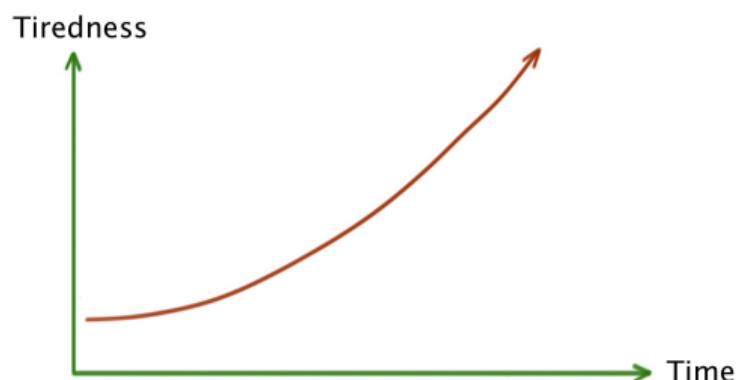
Disturbances are uncontrolled and unmeasured variables

Any potential impact on our system which

- ▶ is not controlled, and
- ▶ is not measured

In the ginger biscuit example, it might include:

- ▶ ambient humidity
- ▶ ambient temperature
- ▶ impurities in the ingredients
- ▶ other examples ...



“Control” and “Measure”: let’s clarify what we mean by those terms
Given enough resources (usually money), we can control and measure most things.

Control

We can not actively control these:

- ▶ ambient humidity (unless indoors)
- ▶ ambient temperature (unless indoors)

Measure

Some variables are too expensive or unreliable to measure.



[Flickr: sidelong]

Uncontrolled and unmeasurable variables can negate all your hard work

Experiment	A	B	C	Time finished
1	—	—	—	08:32
2	+	—	—	09:46
3	—	+	—	10:50
4	+	+	—	12:05
5	—	—	+	13:16
6	+	—	+	14:30
7	—	+	+	15:57
8	+	+	+	17:09

* note the experiments are in *standard order this time*

Take a moment to think about this ...

Pause the video: what problems do you think might be caused by this confounding?

- ▶ Recall what the term “confounding” means (from the previous video)
- ▶ Guess what will happen when you analyze the outcome variable, y
- ▶ Now resume the video ...

when experimenting with computer simulations

The same as regular experiments:

- ▶ you must follow a systematic method
- ▶ don't "play around" with the software: trial-and-error

Many experiments are simulations:

- ▶ bridge/building design
- ▶ chemical factory design
- ▶ improve traffic light timing and queuing
- ▶ test a stock market buy/sell strategy

Different to regular experiments:

1. We can often run computer simulations in parallel
2. Computer experiments (mostly*) are deterministic
 - ▶ i.e. if you repeat the experiments, you get the identical results
 - ▶ this indicates there are no disturbances that affect the outcome
 - ▶ this implies you do not need to randomize the order
 - ▶ or even repeat experiments!

* except those that have a random component

Experiments that could be problematic without randomization

Forums / Participation forums / Ideas for experiments

The best way to learn Italian vocabulary

[Subscribe for email updates.](#)

No tags yet. + Add Tag

The best way to learn Italian vocabulary.

Outcome: number of words learned in Italian

Objective: to maximize the number of words learned in Italian (4 days after the learning day).

Forums / Participation forums / Ideas for experiments

experiment on memorizing 20 digit numbers

[Subscribe for email updates.](#)

No tags yet. + Add Tag

I want to experiment on memorizing a 20 digit number.

objective: to memorize and recall the number

variables: 1. i memorize it by writing it with my hand or not
2. whether i memorize in groups of 4/5.... or all at once

any other suggestions for other variables by anyone??

Forums / Participation forums / Ideas for experiments

Improving my performance in the gym

[Subscribe for email updates.](#)

No tags yet. + Add Tag

I go to the gym at lunch on Monday, Wednesday and Friday mostly regularly. Last wee (round two, as in I completed 7 weeks already). I varied the specific exercises but the ! interested can look up "Westside for Skinny Bastards"; I and my friends are following t and Wednesday are heavy lifting days, while Friday is a heavy volume (many reps) day

The outcome I'd like to affect is "lifting performance." If I had a device that measured p I currently don't, I'm going to use a subjective measure primarily: How good I feel liftin/ power); secondarily I'll use the total amount of weight pushed. I'll plan to increase the ! percentage, and see if the total weight pushed matches it.

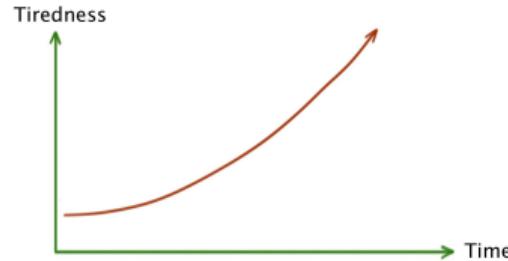
There are a huge number of factors; I'm going to suggest two for the next six weeks:

Other systems experience deterioration over time: we must randomize to counteract this

Gas mileage in a vehicle:



Ginger biscuit baking experiment:



Flickr: mikecogh

(Question: how would you describe the variable: “ $R = \text{rain on the road}$ ”?)

This is a slow moving disturbance: we must account for it though if we run experiments that span several months, or years.



ADVICE

Always randomize!

Why do we randomize?

So when we analyze the effect of the factors – we can be almost certain they are not confound by unmeasured, and uncontrolled disturbances.

ADVICE

disturbances which can be measured, should be recorded

Experiment*	Factors			Measured disturbances		$y = \text{breakability}$
	A	B	C	Temperature	Humidity	
1	—	—	—	23	32	
2	+	—	—	21	56	
3	—	+	—	24	24	
4	+	+	—	22	24	
5	—	—	+	23	30	
6	+	—	+	22	54	
7	—	+	+	23	36	
8	+	+	+	24	24	
9	⋮					

* Experiments were run in random order; but are reported in standard order

“measured disturbances” = “variables which are measured, but not controlled” = “covariates”

Using the covariate information

Experiment*	Factors			Measured disturbances		$y = \text{breakability}$
	A	B	C	Temperature	Humidity	
1	—	—	—	23	32	4
2	+	—	—	21	56	5
3	—	+	—	24	24	5
4	+	+	—	22	24	6
5	—	—	+	23	77	3
6	+	—	+	22	54	8
7	—	+	+	23	36	6
8	+	+	+	24	24	9
9	0	0	0	23	39	5
:						

$$\hat{y} = b_0 + b_A x_A + b_B x_B + b_C x_C + b_{AB} x_A x_B + \cdots + \underbrace{b_T x_T + b_H x_H}_{\text{covariate terms are added to the model*}}$$

* but you will require more than 8 experiments to build this model

Cellphone app example: “CalApp”



Flickr: williamhook

The app has various upgradable features, called “in-app purchases”

- ▶ sync-to-other-devices
- ▶ text message reminders
- ▶ integrate with desktop calendar

Your marketing idea for experimenting:

- ▶ each test group has 2000 people
- ▶ calculate the percentage using the app after 60 days; that's your outcome, y

Read the Harvard Business Review article at <http://yint.org/hbr-article>

Cellphone app example: “CalApp”

The factors you are actively testing (confirm whether they are controllable)

	Low level –	High level +
A: “Promotion”	1 free in-app upgrade	30-day trial of all features
B: “Message”	“CalApp has your schedule available at your fingertips, on any device.”	“CalApp features are configurable; only pay for the features you want.”
C: “Price”	in-app purchase price is 89c	in-app purchase price is 99c

Cellphone app example: test your understanding

Are these “disturbances” (*not measured, not controlled*), or
“covariates” (*measured, but not controlled*), or
“neither of these”:

- ▶ **E**: smartphone user’s age
- ▶ **N**: smartphone user’s gender
- ▶ **S**: smartphone user’s connection speed (e.g. cell, or wifi)
- ▶ **R**: amount of free memory (RAM) on the device
- ▶ **F**: whether the advert/message is delivered via ad network G, or ad network H
- ▶ **D**: if the user’s phone is Android or Apple

Participate in the forums and share your opinion at <http://yint.org/cal-app>