



INFORMS Transactions on Education

Publication details, including instructions for authors and subscription information:
<http://pubsonline.informs.org>

Deming's Quality Experiments Revisited

Tava Olsen,

To cite this article:

Tava Olsen, (2007) Deming's Quality Experiments Revisited. INFORMS Transactions on Education 8(1):37-40.
<https://doi.org/10.1287/ited.8.1.37>

Full terms and conditions of use: <https://pubsonline.informs.org/Publications/Librarians-Portal/PubsOnLine-Terms-and-Conditions>

This article may be used only for the purposes of research, teaching, and/or private study. Commercial use or systematic downloading (by robots or other automatic processes) is prohibited without explicit Publisher approval, unless otherwise noted. For more information, contact permissions@informs.org.

The Publisher does not warrant or guarantee the article's accuracy, completeness, merchantability, fitness for a particular purpose, or non-infringement. Descriptions of, or references to, products or publications, or inclusion of an advertisement in this article, neither constitutes nor implies a guarantee, endorsement, or support of claims made of that product, publication, or service.

Copyright © 2007, INFORMS

Please scroll down for article—it is on subsequent pages



With 12,500 members from nearly 90 countries, INFORMS is the largest international association of operations research (O.R.) and analytics professionals and students. INFORMS provides unique networking and learning opportunities for individual professionals, and organizations of all types and sizes, to better understand and use O.R. and analytics tools and methods to transform strategic visions and achieve better outcomes. For more information on INFORMS, its publications, membership, or meetings visit <http://www.informs.org>

Deming's Quality Experiments Revisited

Tava Olsen

Washington University in St. Louis

Olin Business School

Operations and Manufacturing Management

olsen@wustl.edu

Abstract

This paper revisits two classic experiments used by W. Edwards Deming to teach concepts of quality management. The first, the red bead experiment, demonstrates how managing based on underlying system randomness can be highly problematic. The second, the funnel experiment, demonstrates how adjusting a system based on underlying random variations can significantly degrade system performance. This paper describes adjustments to these experiments that retain their key learning points while making them more suitable for in-class demonstrations and increase the "fun" component of the experiments.

Editor's note: This is a pdf copy of an html document which resides at <http://ite.pubs.informs.org/Vo8No1/Olsen/> (Volume 8, Number 1, September 2007)

1. Introduction

In-class demonstrations, or experiments, are techniques of experiential learning designed to engage the student in the learning process. While their effectiveness as a teaching technique is still being debated (see, e.g., Gosen and Washbush, 2004) there is no question that they can liven up the classroom, engage the students, and make the learning environment more entertaining. This paper describes modifications to two classic quality management experiments with these latter goals, namely increasing student engagement and entertainment, as the primary objectives.

Quality management is one component of many core courses on operations management. However, anecdotal evidence suggests that increasingly instructors are eliminating its inclusion, particularly from MBA core courses, due to students' disinterest in the material. This seems contrary to industry trends such a *six sigma* and certainly contrary to my own personal belief in the importance of the material. One way to ensure that quality remains in the operations management MBA curriculum is to have teaching materials that retain student interest.

W. Edwards Deming, commonly called one of the "gurus of quality", understood that most people do not intuitively understand the underlying principles of quality management; he spent a large portion of his

career trying to effectively teach these principles. He frequently used experiential techniques (see, e.g., Deming, 2000). This paper describes the adaptation of two of Deming's classic experiments, namely, the *Red Bead Experiment* and the *Funnel Experiment*.

The overarching goal of the quality management class where I run these experiments is to teach that effective quality management requires the use of clear analytic thinking and statistical processes. I have chosen the red bead and funnel experiments as particularly effective demonstrations of that principle.

I have taught the modified experiments for five years. I typically teach them to a class of around sixty professional MBA students in their core operations class. Classes in this program are held once a week for three hours in the evening. I devote one of twelve classes to quality management. Given that students have been at work all day, I believe that student engagement and entertainment is particularly important for the learning environment. However, I believe these experiments could be used in any quality management class at either the undergraduate or graduate level.

This paper is organized as follows. Section 2 describes the modifications to the red bead experiment. Section 3 describes the modifications to the funnel experiment. Section 4 concludes the paper.

2. The Red Bead Experiment

In Deming's red bead experiment⁽¹⁾ (see, pp. 346 - 352 of Deming, 2000) the primary learning objective is that "it is all too easy to blame workers for faults that belong to the system." In this experiment, blindfolded "workers" dip paddles into a bin of red and white beads, where red beads correspond to defects. The system is managed according to the production of defects and control charts are taught along the way.

The key takeaways I believe the red bead experiment demonstrates are as follows:

- A system in control will still have variation; in fact, it may have a lot of variation.
- If a system is in control then only management has the power to reduce defects because only management has the power to change the process.
- Pleading with workers or setting quotas will have no effect other than demoralizing the workforce.

In order to teach these takeaways, I have adapted the red bead experiment to meet two further goals:

- Shorten the time required.
- Make the demonstration more fun.

The full red bead demonstration is quite intricate with promotions and demotions based on rate of defect production. I do not personally have time to perform the full experiment in class. Further, it is hard to get students excited about beads; I therefore use red M&Ms® as my source of defects. For those opposed to the use of candy in-class, I am sure other sources could be used, but I find that the statistical variation of number of red M&Ms® in a bag is highly appropriate for the experiment. Further, the M&M® website⁽²⁾ publishes the manufactured proportion of each color in the bag, allowing for an interesting comparison with the sampled values.

The experiment proceeds as follows:

- Hand out a "fun size" bag of milk chocolate M&Ms® to each student in the class (available in bulk packages). For three students hand out a larger 1.69oz counter-side sized bag instead. Try to make this deviation non-obvious by choosing

students who are sitting somewhat separately from the rest of the class.

- Have the students count the number of red M&Ms® and the total number of M&Ms® in their bag. Have them write these two numbers plus their name on a piece of paper and hand the paper in.
- Off-line (my class has a break in the middle but a teaching assistant could be given this duty), enter into a spreadsheet the data with a column for name, a column for number of reds, and a column for total number in the bag.
- Compute an np -chart (see below) for the number of red M&Ms®
- Explain that red M&Ms® represent defects.
- Reward the students who produced the fewest number of defects with an extra bag of M&Ms® and scold the students who produced the greatest number of defects.
- Discuss why this reward mechanism isn't fair and have them provide examples from real life where rewards were based on randomness rather than actual performance. Highlight key takeaways (a) - (c) above.
- Reveal that some students had larger bags. With luck at least one of these bags lies above the control limit (I have never had this fail). Discuss the definitions of normal or common causes and special or assignable causes.

The learning objectives of the modified experiment are as follows:

- Understand and be able to explain key takeaways (a) - (c) above.
- Understand and be able to use np control charts.
- Understand and be able to define what it means for a system to be in control.
- Be able to define and explain normal or common causes and special or assignable causes for variation.

2.1. np -chart Creation

Figure 1 shows a typical np -chart created from the above experiment. This is the actual chart that I created in class, in spring 2003. By placing the pointer over any point on the series, I can highlight the student responsible for that data point (Aaron in the case in Figure 1). Further, while I did indeed hand out three larger bags, only one bag shows a statistically significantly higher number of red M&Ms®. This allowed for a relatively in-depth discussion on the concept of statistical significant.

⁽¹⁾ Deming attributes the original demonstration to Mr. William A. Boller of the Hewlett-Packard Company

⁽²⁾ <http://us.mms.com/us/about/products/milkchocolate/>

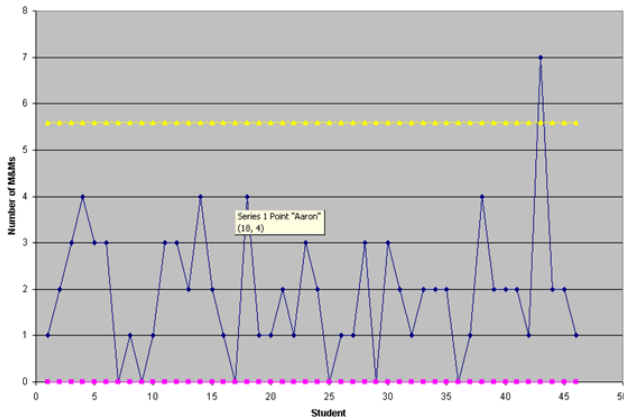


Figure 1: np-chart from in-class data

The upper and lower control limits are shown on the chart and are given by the formulas

$$UCL = \bar{x} + 3\sqrt{\bar{x}(1-\bar{p})}$$

and

$$LCL = \max\left\{0, \bar{x} - 3\sqrt{\bar{x}(1-\bar{p})}\right\},$$

where x and p are the average number and proportion of red M&Ms®, respectively.

From a statistical standpoint, this experiment is not ideal for the creation of np-charts. The data size is limited, and, more importantly, the number of M&Ms® in a bag varies from bag to bag (and certainly to the large bags) making an np-chart, which assumes a constant number in each bag, statistically dubious. A p-chart would be more appropriate, and particularly one designed for variable sample sizes. When I teach the material, I simply point this out, tell the students that there are p-charts available for variable sample sizes, and leave it at that. The more statistically rigorous instructor could either create a variable sample-size p-chart or simply tell the students to only count the first 20 M&Ms® in the bag. In both these cases, one would not expect the larger bags, which have the same *proportion* of red M&Ms®, to be outliers. "Holiday" bags, with differing color proportions, may be an alternative for the instructor trying to retain the production of statistical outliers on the chart. Such outliers are beneficial for the learning objectives because they provide a concrete example of a special or assignable cause; the outliers are caused by a process-shift (shift in bag type).

3. The Funnel Experiment

In Deming's funnel experiment⁽³⁾ (see, pp. 327 - 328 of Deming, 2000) the primary learning objective is "to demonstrate the unbelievable losses from over adjustment." In this experiment, a marble is dropped through a funnel at a target and the point where it lands is marked. The funnel is then adjusted based on where the marble landed relative to the target in four possible ways, namely:

1. No adjustment
2. If the marble lands at point z_k relative to the target then move the funnel by $-z_k$ relative to its current position.
3. If the marble lands at point z_k relative to the target then move the funnel to position $-z_k$ relative to the target
4. Set the funnel directly over where the marble last landed

All of adjustments (2) - (4) lead to poorer performance than no adjustment (1), with adjustments (3) and (4) being unstable. Adjustment (4) will behave like a symmetric random walk. The reason for it's inclusion is that it represents the scenario of trying to match previous performance (e.g., previous color) without reference to the original target.

The key takeaway I believe the funnel experiment demonstrates is as follows:

Variability is present in all processes. Recognizing this and not reacting to or trying to assign causes to inherently random events is the first step in controlling this variability.

In order to teach this takeaway, I demonstrate an adaptation⁽⁴⁾ of the funnel experiment designed to meet four further goals, namely:

1. Make the demonstration more visible to a large audience.
2. Shorten the time required for an in-depth experiment.
3. Make the demonstration more interactive.
4. Make the demonstration more fun.

⁽³⁾ Deming attributes the original demonstration to Dr. Lloyd S. Nelson but also gives other acknowledgements in Deming (2000).

⁽⁴⁾ I learned this adaptation from a former colleague Joseph Milner.

The adapted experiment makes use of a Nerf® dart gun, obtainable at any toy store. It also makes use of Monte Carlo simulation, in this case through Excel. The experiment proceeds as follows:

- i. Draw a vertical line in the middle of the blackboard
- ii. Reveal the dart gun and ask for a student volunteer
- iii. Have the volunteer stand about 15 feet from the board and aim at the line. Mark with an x the point where the dart hits the board.
- iv. Ask the class where the volunteer should aim now. If the target has moved draw a line for the new target.
- v. Repeat steps (iii) and (iv) a couple of times.
- vi. Talk through Deming's funnel experiment and show simulation output (see below) from the four types of adjustments. Discuss the key learnings.

At least some portion of the class will tell the volunteer to aim left if the first dart goes right (and vice versa). There is also usually someone who has done rifle shooting and disagrees. I usually try to hold discussion on this until after step (v) is completed. I also usually bring in sports analogies, such as golf, to emphasize just how bad an idea it is to adjust the target based on one random output.

The advantage of using a vertical line, rather than a fixed target is twofold. First, dart guns tend to be highly inaccurate with respect to height. Second, this significantly simplifies the discussion on where to move the target down to two directions, namely left or right.

In talking through Deming's funnel experiment, I also discuss the four types of adjustments. With luck proponents for adjustments (1) - (3) have been revealed in steps (iii) - (v). Adjustment (4) initially seems counterintuitive to many students. However, it has a natural shooting analogue, namely shooting for grouping, where the goal is to have the bullets within as small a radius as possible and proximity to the bulls-eye is irrelevant. The production application, as mentioned above, is one where the operator is trying to make each unit as close to the prior piece produced as possible.

3.1. Monte Carlo Simulation

I have a spreadsheet⁽⁵⁾ that samples Uniform[0,1] random variables and then adjusts the target according to the four types of adjustment rules given above. Figure 2 presents a sample chart.

I use this chart to show how type 1 is more consistent and stable than the other four types of adjustments. We discuss the chart along with real-life situations that students have seen where one of these adjustment rules was used. We then extend the discussion to a general discussion on management techniques that may hurt rather than help quality.

4. Summary

This paper has presented two adjustments of classic quality management experiments. I have applied these adjusted experiments successfully in my operations management core class for over five years. I find that the in-class nature of the experiments make for much better class discussions than hypothetical examples. Further, both experiments are clearly "fun", which my students appreciate having worked a full day of work before coming to class in the evening. In writing this paper, it is my hope that others will find these adaptations useful, and because of them will be more inclined to teach quality management as part of core operations courses.

References

- Deming, W. E, (2000), *Out of the Crisis*, First MIT Press Edition, MIT Press, Cambridge, MA.
- Gosen, J., and J. Washbush, (2004), "A review of scholarship on assessing experiential learning effectiveness," *Simulation & Gaming*, Vol. 35, No. 2, pp. 270-293.

⁽⁵⁾ The original source for this spreadsheet was my colleague Yossi Aviv.