Using Projects in an Advanced Design of Experiments Class:

A Bubble Mixture Experiment Example

By

Stefan H. Steiner[1], Michael Hamada[2], Bethany J. Giddings White[3], Vadim Kutsyy[4], Sofia Mosesova[3], and Geoffrey Salloum[5]

revised April 21, 2006

Abstract

This article gives an example of how student-conducted experiments can enhance a course in the design of experiments. We will focus on a project whose aim is to find a good mixture of water, soap and glycerin for making soap bubbles. This project is relatively straightforward to implement and understand. At its most basic level the project introduces students to mixture experiments and general issues in experimental design such as choosing and measuring an appropriate response, selecting a design, the effect of using repeats versus replicates, model building, making predictions, etc. To accommodate more advanced students, the project can be easily enhanced to draw on various areas of statistics, such as generalized linear models, robust design, and optimal design. Therefore it is ideal for a graduate level course as it encourages students to look beyond the basics presented in class.

Key Words: mixture experiments, design of experiments, robust parameter design, generalized linear model, poisson regression, optimal design, constrained experimental region

1 Associate Professor in the Department of Statistics at the University of Waterloo, ON, Canada.
2 Technical staff member in Statistical Sciences at Los Alamos National Laboratory, Los Alamos, NM, USA.
3 Graduate student in the Department of Statistics at the University of Waterloo, ON, Canada.
4 Cytokinetics, Inc., San Francisco, CA, USA.
5 Department of Mathematics, Camosun College, Victoria, B.C. Canada.

Introduction

Effective instructors, in addition to presenting course material, challenge students to think critically about what they are studying and to consider possible applications of the material. One way to accomplish this is to assign projects specifically targeted to address material covered and to encourage students to think beyond what they have learned in class. This is especially important in graduate level courses as compared to those at the undergraduate level, as the focus shifts toward more independent thinking. This can be a difficult task, especially when some graduate students have taken the equivalent undergraduate course, or when the course is cross-listed as graduate and undergraduate. In the latter case, an instructor has to accommodate both sets of students in the classroom, challenging them simultaneously. One way of handling this is to assign graduate students a project of broader scope. Another option is to assign the same project to both undergraduate and graduate students. The undergraduates would be required to do a simpler experiment that only requires a basic analysis, whereas graduate students would be expected to expand upon the ideas. In this paper we will describe one such project that was implemented in a cross-listed undergraduate and graduate course in Design of Experiments. It involves the application of many basic concepts in Experimental Design, but also extends beyond what is typically taught in such a course.

The experiment involved the simple task of blowing soap bubbles. The object was to determine an optimal bubble solution consisting of a mixture of water, dishwashing soap and glycerin. Minimal prior knowledge of the science behind bubbles is necessary to complete this project. However, in addition to the educational value associated with planning, conducting and analyzing an experiment, students also have the opportunity to learn some science; bubbles are soap films which people have devoted their entire careers to studying (Boys, 1959; Isenberg, 1992). Appendix A contains some background on the science behind bubbles.

In our teaching experience three separate groups of students have conducted versions of this project. The experimental design and data come from one of these groups. For the illustrative purposes of this article, we present a possible analysis of these data. The instructor has the flexibility to vary what is required and the amount of guidance offered depending on the level of the students' knowledge. Throughout the paper, we will discuss how this project can be adjusted for use in teaching, as there are many different ways that the bubble problem can be addressed.

We present the issues and choices that the students need to confront in completing this project by following the structured empirical problem-solving framework PPDAC (Oldford and MacKay, 2001). PPDAC is an acronym for a five-stage process:

>   **P**roblem: Develop a clear statement of what we are trying to learn
>   **P**lan: Determine how we will carry out the investigation
>   **D**ata: Collect the data according to the plan
>   **A**nalysis: Analyze the data to answer the questions posed
>   **C**onclusion: Draw conclusions about what has been learned

We find PPDAC useful to structure the design and analysis of any experiment and have used it in our teaching. We present the bubble mixture experiment project by discussing each of the five stages of PPDAC. At each stage, we describe the choices made in our example project along with their rationale. As well, we discuss options for making the project either less or more complicated.

However, we start by giving some background on mixture experiments. To complete this project students must recognize this as a mixture experiment and apply the appropriate methods. This may mean independent research for the students.

Mixture Experiments

In the proposed experiment we model the effect of the proportions of the three components (denoted W for water, S for soap and G for glycerin). Such experiments are known as mixture experiments (Cornell, 2002). In a mixture experiment, each proportion may vary between zero and one, and the components must sum to one. As a result, with three components, as in our example, the experimental region is a triangle, defined by the points (0, 0, 1), (0, 1, 0) and (1, 0, 0). Traditionally, this experimental region has been viewed in two dimensions as displayed in Figure 1 that employs a triangular simplex coordinate system. Each point lying on or in the interior of the triangle represents a possible mixture. The plot can be interpreted in the following way. Each vertex of the triangle represents a pure mixture (100% of one component and none of the rest). As one moves away from this vertex toward the opposite side, the proportion of the component gradually decreases to zero. Points interior to the triangle have some proportion of each of the three components. For example, the point shown by the black dot in Figure 1 corresponds to a mixture of 80% water, 15% soap and 5% glycerin, i.e., (W, S, G) = (0.8, 0.15, 0.05). The project could be extended to include more components but that would increase complexity, especially in terms of the graphical displays.
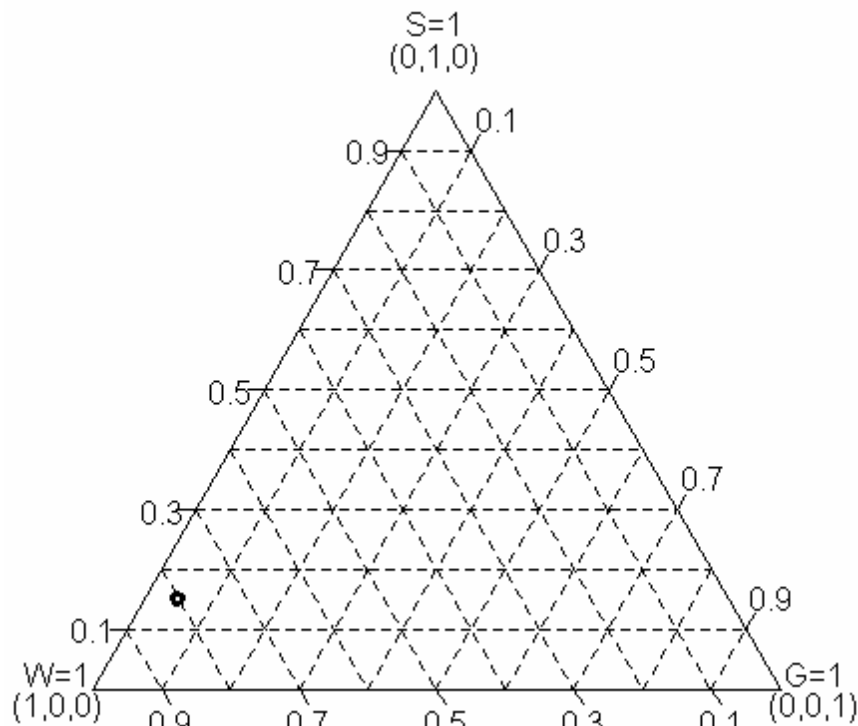

Figure 1: A Two Dimensional View of the Three Component Mixture Experimental Region
All mixtures must lie on or inside the triangle whose equation is W+S+G=1.

In this application, the experimental region is further constrained because we know, before running the experiment, that some combinations will not produce any bubbles (e.g., 100% water) or be very poor. We discuss this issue further in the Plan section.

For a mixture experiment, the standard linear first-order response model for a normally distributed response Y is

$$Y \sim N(\mu, \sigma^2), \text{ where } \mu = \beta_W W + \beta_S S + \beta_G G. \tag{1}$$

Note that there is no intercept in this model. If the intercept were included, the design matrix would be singular because the proportions of water, glycerin and soap must add to one. Additional models commonly used for mixture experiments include the quadratic model, special cubic model, and the rarely used full-cubic model (Cornell, 2002):

Quadratic:
$$\mu = \beta_W W + \beta_S S + \beta_G G + \beta_{WS} WS + \beta_{WG} WG + \beta_{SG} SG \tag{2}$$

Special-Cubic:
$$\mu = \beta_W W + \beta_S S + \beta_G G + \beta_{WS} WS + \beta_{WG} WG + \beta_{SG} SG + \beta_{WSG} WSG \tag{3}$$

Full-Cubic:
$$\mu = \beta_W W + \beta_S S + \beta_G G + \beta_{WS} WS + \beta_{WG} WG + \beta_{SG} SG$$
$$+ \alpha_{WS} WS(W - S) + \alpha_{WG} WG(W - G) + \alpha_{SG} SG(S - G) + \beta_{WSG} WSG \tag{4}$$

Problem

In the Problem stage, the students must clarify their objective by defining what they mean by an optimal bubble solution. To do this they need to choose a target population, that is the conditions over which they would like their conclusions to be valid, and select an appropriate response variate and population attribute.

The choice of response and population attribute is tied to the definition of best or optimal bubble mixture, and has a big impact on an appropriate analysis of the subsequent experimental data. In our example, the students decided an optimal solution was one that, on average, produced the most bubbles. Therefore, the response of interest was the number of bubbles and the attribute was the *average* number of bubbles across all conditions in the population. Other possible choices for response include: the size of the biggest bubble, the time the bubbles survive, whether or not any bubbles where made, etc. In each case, it is critical that the student think about both whether the response reflects their definition of optimal, and whether the proposed response can be precisely and accurately measured. Students can be quite inventive; for instance, to measure time, they thought of blowing bubbles onto a dish so that the bubble can be observed until it pops. A possible extension to the basic project is to use more than one response. This will lead to multiple analyses and possibly the need to compromise in some way when making recommendations.

In the choice of a response, the instructor needs to be careful the project does not become too complex for the students to handle. To assess this, we need to think ahead to the analysis stage. For instance, with a response defined as the number of bubbles, a standard regression model with normally distributed errors may need to be replaced by a Poisson regression model. Poisson regression is an advanced topic that students would learn about in a course on generalized linear models. Alternatively, a standard analysis with the usual normal error assumption that ignores the discrete nature of the count response may be acceptable. An appropriate choice depends on the data. Another possibility is an analysis on a transformed scale. For instance, in another application of this experiment, time to a bubble popping was used as the response and the log time was well modeled using a standard normal regression model. The instructor should be prepared to discuss the pros and cons of each modeling choice. The choice of response can be a great lead in to a class discussion of model assumptions and more advanced types of models.

Next, the students needed to define their target population by answering the question "average over what?" For example, they may decide they want to find the best bubble solution for all different conditions, including different bubble making devices, environmental conditions, etc. Alternatively, they may decide to look for the best bubble solution only for typical summer conditions in their part of the world. In a class project, the students have control over how ambitious they want to be. In other applications, the goals of an experiment are usually driven by outside considerations, and we need to try to achieve them through a good plan that makes appropriate use of the experimental principles of blocking, replication and random assignment. The students typically come back to the definition of the population after they have given more thought in the Plan stage to what is realistic, given the time and cost constraints inherent in a class project.

Plan

In the Plan stage the students must determine the details of the experiment they will run to try to find the optimal bubble solution as defined in the Problem stage. More specifically, they need to select the experimental factors and levels, choose a design and run order, worry about measurement of the response and think about the logistics of the data collection.

In our experience, a cause and effect diagram (Ishikawa, 1982) is a useful tool in the Planning stage. To produce a cause and effect diagram, also called a fishbone or Ishikawa diagram, we attempt to list all possible important explanatory variates that may influence the response. This is typically done using a team of subject matter experts, but for the bubble mixture experiment most people's childhood experience and common sense are sufficient. In the diagram, to help organize our thoughts, we use major branches to collect related explanatory variates. An example of a possible cause and effect diagram for the bubble mixture experiment is shown in Figure 2.
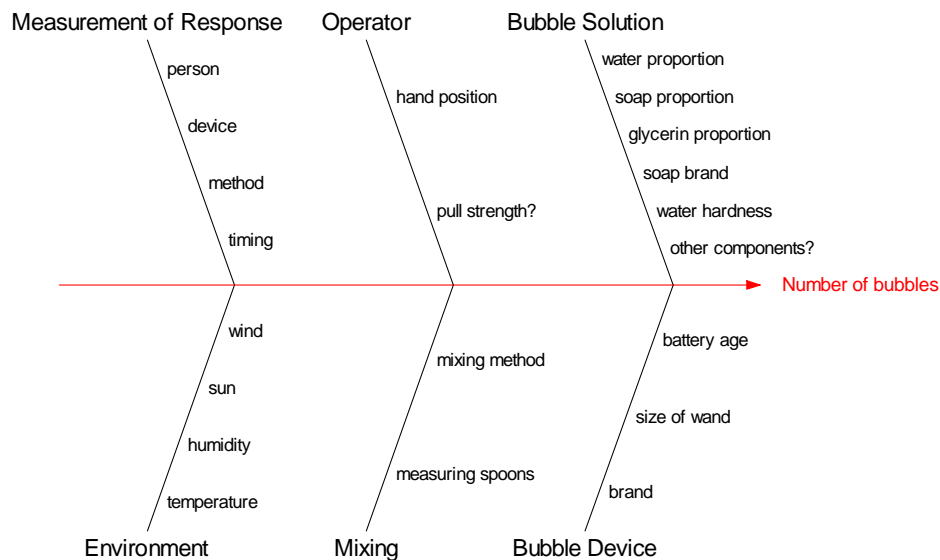
Figure 2: Bubble Experiment Cause and Effect Diagram

While the students should not spend a large amount of effort developing a cause and effect diagram, they should try to think broadly about all the possible inputs that could influence their chosen response. There are many explanatory variates that could (with our limited knowledge) influence the number of bubbles produced, of which many are not mixture components. Consideration of all the possibly important explanatory variates can have a large impact on the experimental plan.

In the example project, the students decided to deliberately vary the water type and soap type in the experiment as well as the mixture components water, soap and glycerin. By choosing two different water types, spring and tap, and two different soap brands, joy and ivory, they planned to also consider whether the optimal bubble formulation depended on the levels of these two noise factors. As well they decided to look for a robust bubble solution that works well no matter what sort of water or soap is used. See Taguchi (1987) for more on robust design experiments. Using the robust design terminology we shall refer to soap and water type as noise factors.

What should be done with all the other possibly important explanatory variates? One option is to (try to) hold them fixed during the experiment. This makes the experiment easier to run (assuming it is not difficult to hold the explanatory variates fixed), but restricts the generalizability of the conclusions. Remember the choice of target population in the Problem stage.

Another option is to create blocks within which some explanatory variates are held fixed, and then replicate the design over a number of blocks. For instance, we could block by day by repeating the design over two (or more) different days, if we thought there might be a large day effect. This helps achieve the goal of seeing the effect of changing the mixture components while at the same time allowing us to check that the results obtained under one set of conditions (one day) are also valid for other conditions (i.e., the second day).

In the example the students decided to hold all other explanatory variates fixed. This choice was made mostly because it was easy. Blocking was rejected because the planned experiment was already complex with three mixture factors and two noise factors. Also, while they felt that temperature and humidity were probably important explanatory variates, running an experiment where they changed the temperature and humidity was not practical given the available resources.

To accurately compare the bubble formulations, they used a consistent device to blow bubbles. A battery-operated device avoided problems associated with a person blowing bubbles such as speed of a person's breath and temperature. An assortment of bubble-blowing toys were considered and tested with a standard store-bought bubble solution. One of the most consistent battery-operated devices was chosen to run the experiment. To ensure reliable results, the batteries were replaced prior to running the experiment. Also for consistency, each student had one task in the running the experiment. One student made all the bubble solutions; a second operated the bubble making device, while the third student did all the bubble counting and recording. This was a good idea as it reduced the chance that student-to-student differences could influence the results.

The cause and effect diagram may suggest we need to further clarify what we mean by an optimal bubble mixture. For example, we may have originally decided to try to find the best bubble mixture for a wide variety of conditions. However, based on limited resources and time, running an experiment that would reasonably allow us to draw such general conclusions may not be feasible. As a result, as in the example, we may decide to look for the optimal bubble mixture with a restriction on some of the explanatory variates.

Standard experimental designs for estimating all the parameters in the mixture experiment models given in an earlier section are outlined in Cornell (2002). However, for the proposed experiment, a constraint on the experimental region makes sense. The students reviewed some bubble solution recipes found in children's books, the Internet, teaching materials and toy manufacturers' suggestions. The recommendations were similar; all restricted the proportion of glycerin to less than 0.15 and the

proportion of the dishwashing soap to less than 0.35. At the same time, the proportion of water should not exceed about 0.98 since 100% water will not produce bubbles.

These considerations led the students to the following constraints on the experimental region: $G \leq 0.15$, $S \leq 0.35$, and $W \leq 0.98$; the first two constraints imply that $0.50 \leq W$. This selected constrained experimental region is displayed in Figure 3. These decisions are somewhat arbitrary and should be reviewed for future projects using the results presented later in this article. We are not aware of any readily available software that can plot a restricted simplex region. Having the students create these plots in R (R Core Development Team, 2004) is a valuable learning experience in itself. As we will see in the Analysis section, it will be useful for generating contour plots of the predicted response.
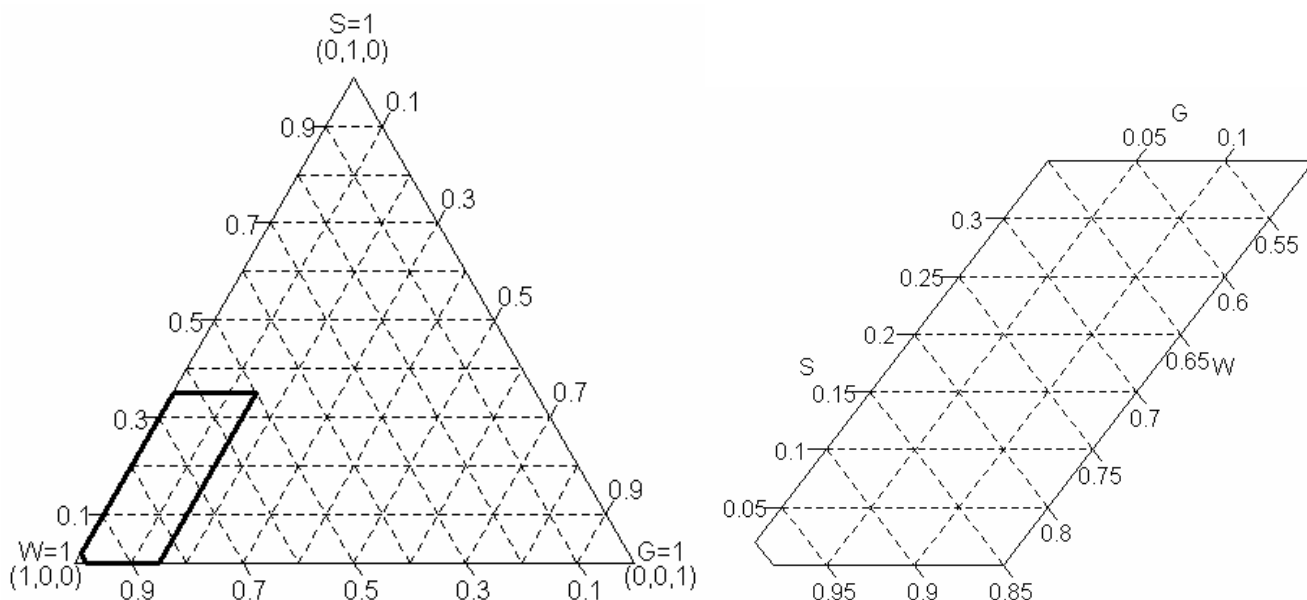

Figure 3: Constrained Experimental Region $G \leq 0.15$, $S \leq 0.35$, $0.5 \leq W \leq 0.98$
Left panel shows constrained region relative to complete simplex.
Right panel shows the constrained region in close up.

Additional considerations could be included in defining the experimental region. Because glycerin is much more expensive than dish soap which is in turn more expensive than water, further constraints could be added to informally consider cost. For example, we may decide to restrict water to greater than 60% (i.e., $W \geq 0.6$). Alternatively, the cost constraints could be imposed more formally through a cost function.

From the restricted experimental region the students must select a design. In the example, the students selected 12 component mixtures spread out over the constrained experimental region based on their judgment. The 12 chosen mixtures are shown in Figure 4. They used 12 formulations because they felt that was a manageable number and would allow them to fit the special cubic model (3) while still retaining some degrees of freedom for assessing model fit. The chosen design does not cover the constrained region well in places. In particular, based on some trial and error, the students decided they also needed a lower limit on the amount of soap. As a result, they only selected designs with $S \geq 0.04$. This, in effect, further restricts the design region. They also avoided the design point $(W, S, G) = (0.5, 0.35, 0.15)$ that would probably be included in most definitions of an optimal design.
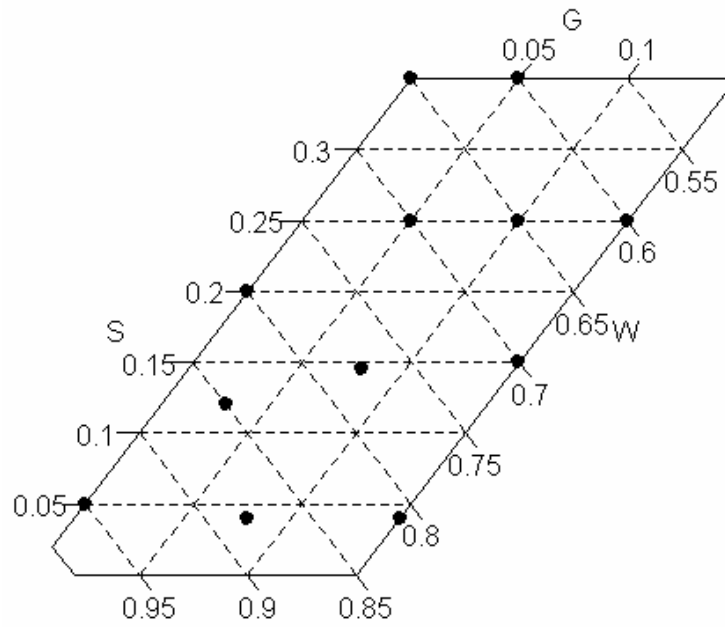
Figure 4: Proposed Mixture Design Showing the 12 Mixture Proportions

Clearly other choices of design points and the number of mixture combinations are possible and are probably better. Students can extend the project by looking for an improved design. However, finding an optimal design within a constrained mixture region is nontrivial. The best design will depend on the proposed model. One possibility for finding near optimal designs is to use a genetic algorithm as in Goldfarb et al. (2005). Finding an optimal design in the context of a generalized linear model makes the task harder still.

To explore how varying the types of soap and water would affect the number of bubbles produced, a crossed control-by-noise array was used. In other words, forty-eight trials were run using all possible combinations of the twelve bubble formulations, two types of water and two types of soap. As mentioned earlier all other explanatory variates in the cause and effect diagram were to be held fixed during the experiment.

Another option would be to add more noise factors and conduct a screening experiment to determine the set of factors that contribute to the effectiveness of the bubble solution. With many noise factors, a fractional design in the noise factors could have been used.

To reduce the effect of potentially unknown confounding variables, the order in which the students created and ran the formulation was randomized. The design and run order (as well as the results) are given in Appendix B.

To partially counteract any variability that could be attributed to the bubble device, the students also planned to pass each solution through the bubble blowing process five times, i.e., they conducted five repeats for each run. The difference between replicates, where each of the 48 bubble formulations would be made again from scratch, and repeats, where the same bubble formulation is used to generate additional responses, is very important for the students to understand. Clearly replication captures additional sources of error, such as mixing variation. We discuss the issue of repeats versus replicates further in the Analysis section.

In some applications of this project censoring of the response could become an issue. If there were a large number of bubbles, it would not be possible to accurately count the actual number of bubbles. Instead, we could then record a right-censored value, e,g., there were at least 25 bubbles. Similarly, with time to a bubble popping as the response, right censoring may arise if the times are too long and the bubbles have not popped after, say, half an hour.

Data

In the Data stage we run the planned experiment and collect the data, noting anything that deviates from the plan.

To run the planned experiment, the students mixed forty-eight solutions of water, soap and glycerin with the appropriate water and soap type, each to a combined volume of 100 ml. They ran the experiment in the same order in which they created the solutions. All runs were done in the office of one of the students over a single afternoon. Thus, environmental inputs such as temperature and humidity were held (approximately) constant over the course of the experiment. For each repeat, the response, that is the number of bubble produced, was determined. Appendix B gives the numbers of bubbles recorded as a result of the experiment.

There were a number of practical issues that arose while conducting the experiment. One of the primary difficulties was measuring the exact proportions of glycerin and soap. Glycerin, in particular, is very viscous and is difficult to pour into solutions. To aid in the process, the students measured the required amount of water first and then tried to remove any remaining glycerin and soap by dipping the measuring spoons into the water. This was a very effective technique, as they seemed to get most of the glycerin and soap off the measuring spoons.

The experiment was time consuming to run, with mixing the 48 solutions taking the bulk of the time. This is clearly an important consideration that should be considered in the Plan stage. Also, it can be very helpful to make a few trial runs before conducting the whole experiment. This allows everyone to get a better idea of the process and time involved and learn a lot about the logistics of running the experiment. If necessary, the experimental plan could be altered and/or scaled back if time constraints appear to be a problem.

Analysis

Before doing any formal modeling, it is always a good idea to examine the data graphically. We start by plotting the five repeat counts by the average count for each of the 48 runs – see Figure 5.
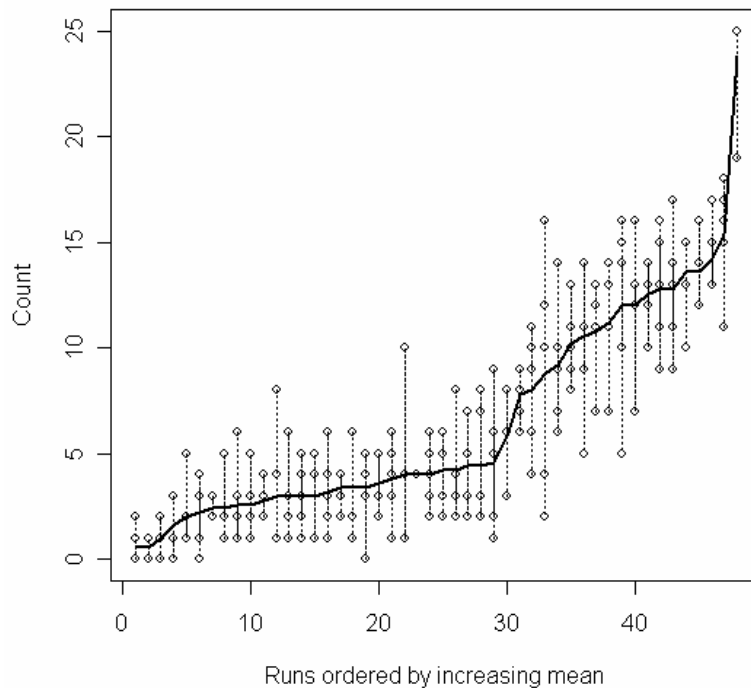
Figure 5: Plot of Number of Bubbles in Each Repeat by Run
Solid line shows the average count at each run.

Figure 5 shows there are no wild outliers. Also, while it is difficult to assess within run (i.e. repeat to repeat) variation from only five repeats, there is no evidence that a Poisson assumption, i.e. that the standard deviation should increase as the square root of the mean, is unreasonable.

In our example there are five repeats per run. There is no replication because each of the 48 bubble solutions was made only once (this was enough work!). In modeling, using the repeats as replicates does not result in an appropriate analysis as the experimental error would likely be underestimated and there would be too many significant effects. The standard approach to handle repeats is to calculate a performance measure, such as an average or sum, etc., across the repeats for each run and to do the analysis using the performance measure. In the example, the students decided to use the total number of bubbles produced across the five repeats as the performance measure. This is appropriate in this case as they planned to fit a Poisson regression and because the sum of independent Poisson counts is still Poisson with mean equal to the sum of individual means.

Next, the students plotted the data to informally examine the effect of the noise factors. Figure 6 shows the total number of bubbles across the five repeats by mixture formulation with different plotting symbols for the two levels of each of the two noise factors. We see clearly that the number of bubbles produced does not seem to depend much on the water type, but depends strongly on the type of soap, with Joy producing many more bubbles on average than Ivory.
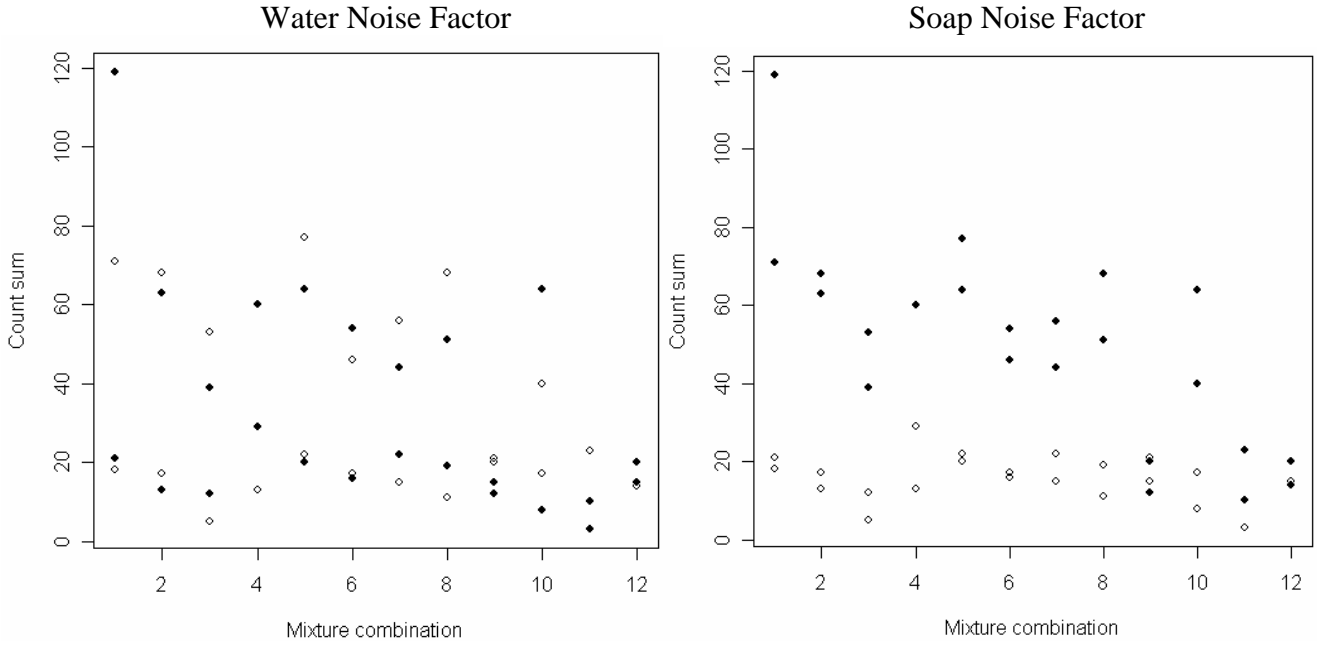
Figure 6: Total Bubble Count (across the five repeats) Versus Mixture Combination
Left plot: Solid circle = Spring, Open circle = Tap for Water noise factor.
Right plot: Solid circle = Joy, Open circle = Ivory for Soap noise factor.

Now they were ready to consider formal models. In the example, there are a number of complications that require extensions to models (1) – (4). More complex models arise with the inclusion of additional covariates (noise factors) that are not a part of the mixture. In the experiment considered in this paper, information was collected on two additional covariates: water type (N_W) and soap type (N_S). The two levels for each covariate were coded –1 and +1. An extension to (1) that includes these covariates is:

$$\mu = \beta_W W + \beta_S S + \beta_G G + \left(\beta_{WN_W} W + \beta_{SN_W} S + \beta_{GN_W} G\right)N_W$$
$$+ \left(\beta_{WN_S} W + \beta_{SN_S} S + \beta_{GN_S} G\right)N_S$$
$$+ \left(\beta_{WN_W N_S} W + \beta_{SN_W N_S} S + \beta_{GN_W N_S} G\right)N_W N_S \qquad (5)$$

Models (2), (3), and (4) can be extended in a similar manner. Note that as we add noise factors, the number of model parameters multiply by the number of noise combinations.

In the example, since the chosen performance measure was the total number of bubbles produced across the five repeats for each run, the students fit a generalized linear model (McCullagh and Nelder, 1989) with a Poisson response and a log link to the data.

In general, the Poisson regression for a response Y is written:

$$Y \sim \text{Poisson}(\theta), \text{ where } \log(\theta) = \beta^T X,$$

and θ is the mean number of bubbles, X is a vector of covariates and β is a vector with the corresponding coefficients. In our example we would model $\beta^T X = \mu$, where μ is given by a model similar to (5).

11

Putting this all together, the students fit both the extended quadratic and special cubic forms of the mixture model. The special cubic model with separate terms for each of the four combinations of the noise factors had 28 parameters while the corresponding quadratic model had 24 parameters. The AIC for the special cubic model was 390.91 and the AIC for the quadratic model was 389.94. Using AIC for model selection, since the quadratic model had a smaller AIC, the students concluded that it was preferred; i.e., the additional cubic terms did not help the model fit.

In the quadratic model fit, the students then noticed that none of the terms involving the water type noise factor were significant. As a result, all such terms were removed from the model leading to what we will refer to as the reduced quadratic model. See Appendix C for the details of the reduced quadratic model fit. There were 5 significant terms: W, WS, GN_w, WGN_w and SGN_w. The model fit is summarized as

$$\log(\hat{\theta}) = 2.4W - 3.4S + 17.8G + 14WS - 17.6WG - 5.3SG$$
$$- 0.2WN_w + 2.3SN_w + 47.3GN_w - 5.2WSN_w - 53.2WGN_w - 69.7SGN_w \qquad (6)$$

The deviance residuals from the reduced quadratic model are plotted in Figure 7. There are no obvious problems with the model fit, and there is no pattern over time.
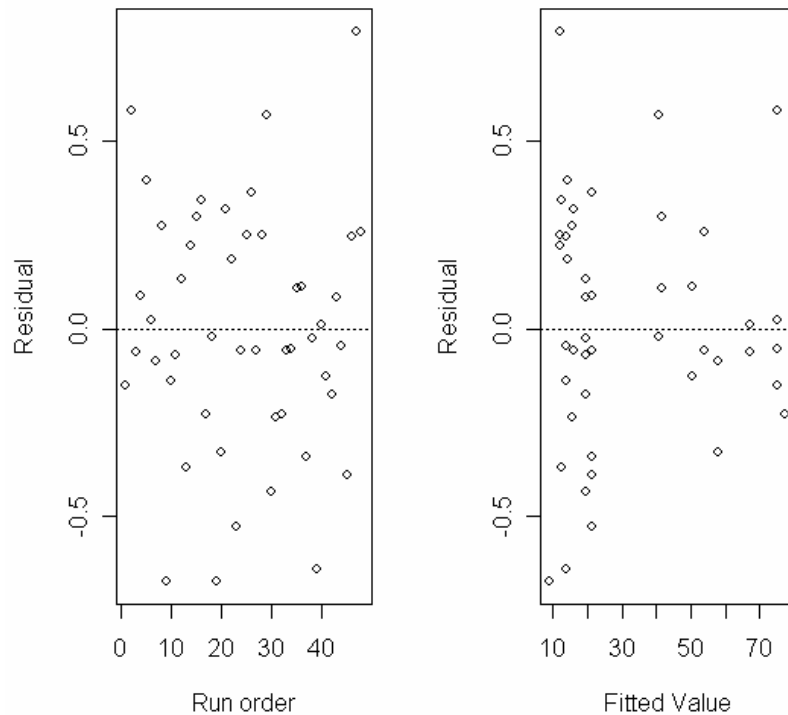
Figure 7: Deviance Residuals From Reduced Quadratic Model Versus Run Order and Fitted Values

To explore the results from the reduced quadratic model fit, contour plots of the response for each of the soap types are given in Figure 8. Note that the response surface gives predictions on a single repeat basis, i.e., we exponentiate (6) and divide by five.

Creating these contour plots over a constrained mixture region was non trivial as there is no existing commercial or shareware software that can perform this task. Developing the necessary code in R was another good learning opportunity for the students. Figure 8 was generated by randomly generating

many points in the constrained region, calculating their predicted response using (6), identifying those points whose response is near some selected values and finally plot the points.
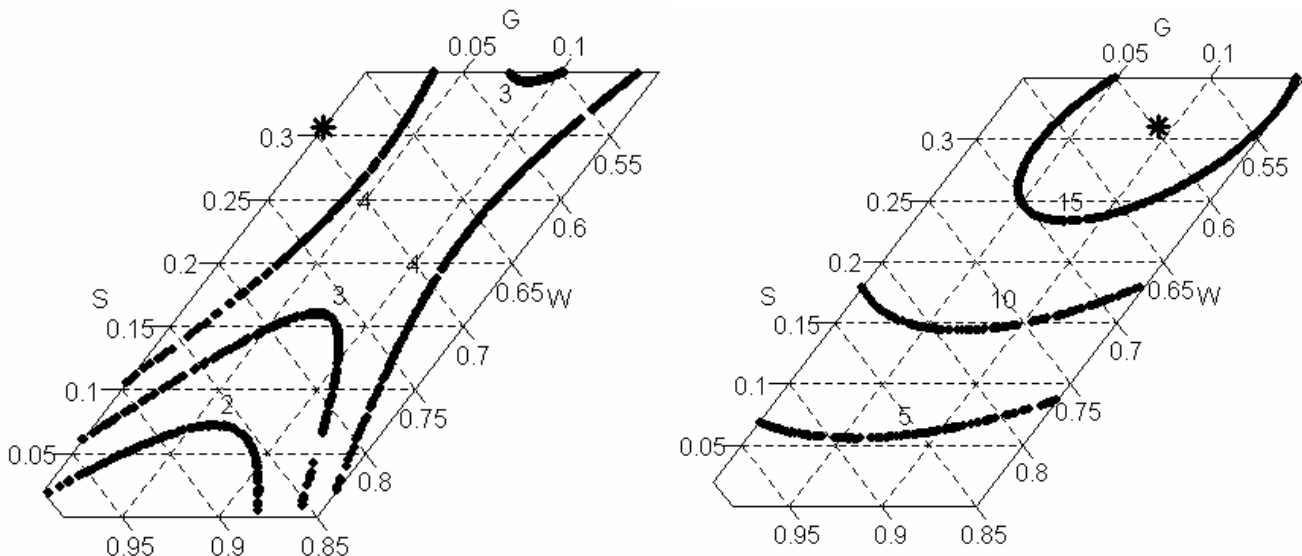
Figure 8: Contour Plots Showing the Predicted Number of Bubbles Per Repeat
Left panel: Ivory soap, Right panel: Joy soap
Asterisk in plot gives the maximum prediction

Based on Figure 8, we can easily see approximately where the best bubble solution is located for each of the two types of soap. Given the fitted model (6), it is also possible to optimize the response more formally using a constrained optimization routine. However, in this case the students obtained the maximums from the contour plot routine. Rounding off, for Ivory soap, the maximum predicted number of bubbles per repeat was 4.3 at the component proportions (W, S, G) = (0.69, 0.31, 0). While for Joy, the maximum was much higher at 16.8, and occurred at the component proportions (W, S, G) = (0.60, 0.31, 0.09). These optimal bubble solutions are shown in Figure 8 using an asterisk.

In the example, the students also looked for a bubble solution that is robust across the two different soap types. The best way to do this is not clear. Looking at the two contour plots in Figure 8 we can use judgment to compromise. For instance, component proportions near (W, S, G) = (0.65, 0.3, 0.05) are predicted to give, on average, close to 15 bubbles for Joy and 4 for Ivory. There are many other possible ways to compromise. For instance, we may instead try to maximize the minimum average response. Another alternative is to use formal optimization, perhaps to minimize the larger-the-better loss, $\text{Loss} = 2\sqrt{\text{Var}_N(y)} - E_N(y)$ (Steiner and Hamada, 1997, equation (5)), where $E_N(y)$ and $\text{Var}_N(y)$ are the mean and variance of the response over the noise factor N = soap type. However, this criterion was developed with the normal regression model in mind. For this experiment, minimizing this loss leads to a solution that drives the variance down at the expense of the mean, and results in a undesirable bubble solution with small variability but also a small mean for both type of soap. An interesting study for the students would be to develop a more meaningful criterion for the Poisson regression model.

Note that in cases where a continuous normal response is chosen, the standard modeling could be used in place of the generalized linear model. Also, without the noise factors there would be a quarter as many runs and no need to consider robustness.

As another alternative to using generalized linear models, the students could do a Bayesian analysis with the Poisson regression model (Gelman et al. 2004). This is relatively easy to do with WINBUGS (Spiegelhalter et. al. 2004). An advantage of a Bayesian approach is that the correlation of the parameter estimates (i.e., regression coefficients) are captured in the draws from the joint posterior distribution of the parameters. Moreover, the uncertainty on predicted means or predictons that are functions of the parameters are easily obtained.

Conclusions

Depending on the original goal the conclusions can be presented in a variety of ways. Much depends on how we view the process (non mixture) variables: water and soap type. One option is to make recommendations for the best levels of the process variables and the associated mixture components. For instances, based on the fitted models the best bubble mixture would result from using joy soap with proportions (W, S, G) = (0.60, 0.31, 0.09). This mixture is predicted to produce on average 16.8 bubbles, and using the Poisson distribution, 95% of the time there should be between 9 and 25 bubbles.

Another option is to give conclusion based on a robust bubble solution that works reasonable well no matter what soap is used. Clearly something in between is also possible. Perhaps we are willing to make a recommendation for the soap type since that can be easily bought, but want a bubble mixture that will work well over a variety of water hardness. In our example this does not change our conclusions since water type was found to have little influence.

In any case, the students should comment on possible limitations to their conclusions. In the example, all environmental variates were held fixed during the experiment. This may cause concern about generalizing the results. Perhaps, for other environmental conditions, a different bubble mixture would be optimal. This seems likely given our belief that temperature and humidity have a strong influence. How much our results are affected by measurement or mixing errors is also a worry. Note also that, in any case, with only two types of dish soap in the experiment any proposed "robust" solution is only robust across Joy and Ivory, not (necessarily) across all brands of dish soap available.

As part of the conclusion, the students could also comment on the choices made in the problem or plan stage. For instance, setting the goal to maximize the average number of bubbles may not have been the best choice if the proposed best bubble solution only produced many very small bubbles. This would probably not meet with the approval of any children testing our proposed solution. In the example, the students concluded that spring and tap water were not the best choices for their two water types. They had intended on comparing two water types with very different mineral content. However, it is likely that both the tap and spring water were quite hard, i.e., had substantial mineral content. This could have been the reason the quadratic model suggested water type was not important. A better choice may have been to use distilled water as one of the two water types.

To strengthen their conclusions, very eager students can conduct an additional simple experiment to confirm that the proposed optimal solution produces close to the predicted number of bubbles on average. If the results can be validated in this way, we would have more confidence in the fitted model and their conclusions in general. Another possibility is to suggest this as further work (for the next students!).

To make the project even larger in scope, the presented experiment could be seen as the first in a series of experiments to look for a better bubble solution. In other words, the students or class could be asked to run a series of experiments and use response surface methods. Though it is not clear how best to do

this in the context of a crossed mixture by noise array. Each new experiment results as another application of PPDAC and would be expected to build on the knowledge gained in previous experiments. For example, each new experiment could further restrict (or move) the constrained experimental region.

Summary

As stated in the Introduction, the goal of this paper is to provide instructors of advanced design of experiments course with a versatile teaching tool. Inherent in any applied project of this kind is the fact that it solidifies the concepts and ideas presented in class for the students. One of the advantages of this particular project is that its level of complexity can be adjusted based on the students background. The physical experiment is easy to conduct and there is minimal time and cost involved. At the same time, students have the opportunity to gain valuable insight into various topics in statistics, such as:

- Experimental Design
- Optimal Design
- Constrained Experimental Regions
- Mixture Experiment Models
- Generalized Linear Models
- Robust Optimization
- Graphical Techniques

In addition to these areas, students are also exposed to group collaboration and obtain valuable experience in report writing and/or presenting. In the project outlined in this paper, a number of statistical/mathematical software packages were used. Although it is not required that students use these particular programs, the experience gained through use of the computer will greatly contribute to their learning. Throughout this paper we made a number of suggestions regarding specific issues that may benefit students at the planning of the experiment in the future. More ideas for projects involving mixture models can be found in Sahrman, Piepel and Cornell (1987), Anderson (1997) and Piepel and Piepel (2000). Further examples in experimental design in general are given in Hunter (1977).

References

Anderson, M.J. (1997), "Bigger and Better Bubbles via Mixture Design," *Stat Teaser*, September, 1 and 3.

Boys, C.V. (1959), *Soap Bubbles Their Colors and Forces Which Mold Them*, New York: Dover Publications, Inc.

Cornell, J. (2002), *Experiments With Mixtures*, third edition, New York: John Wiley and Sons.

Goldfarb, H.B., Borror, C.M., Montgomery, D.C. and Anderson-Cook, C.M. (2005), "Using Genetic Algorithms to Generate Mixture-Process Experiments Designs Involving Control and Noise Variables," *Journal of Quality Technology*, 37, 60-74.

Gelman, A., Carlin, J.B., Stern, H.S. and Rubin, D.B. (2004), *Bayesian Data Analysis*, second edition, Boca Raton: Chapman and Hall/CRC.

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

Isenberg, C. (1992), *The Science of Soap Films and Soap Bubbles*, New York: Dover Publications, Inc.

Ishikawa, K. (1982), *Guide to Quality Control,* second revised edition, Tokyo: Asian Productivity Organization.

McCullagh, P. and Nelder, J. (1989), *Generalized Linear Models*, second edition, London: Chapman and Hall.

Montgomery, D. C. (2001), *Design and Analysis of Experiments*, fifth edition, New York: John Wiley and Sons, Inc.

Oldford R.W. and MacKay R.J. (2001), Stat 231 Course Notes, Fall 2001, University of Waterloo, Waterloo, Ontario.

Piepel, M.G. and Piepel, G.F. (2000), "How Soil Composition Affects Density and Water Capacity - A Science Fair Project Using Mixture Experiment Methods," *STATS*, 28, 14-20.

R Development Core Team (2004), *R: a language and environment for statistical computing*, Vienna: R Foundation for Statistical Computing. {http://www.R-project.org}

Sahrmann, H.F., Piepel, G.F. and Cornell, J.A. (1987), "In Search of the Optimum Harvey Wallbanger Recipe via Mixture Experiment Techniques," *American Statistician*, 41, 190-194.

Spiegelhalter, D., Thomas, A., Best, N. and Lunn, D. (2004), WinBUGS Version 1.4.1 User Manual. (http://www.mrc-bsu.cam.ac.uk/bugs/)

Steiner, S.H. and Hamada, M. (1997), "Making Mixtures Robust to Noise and Mixing Measurement Errors," *Journal of Quality Technology*, 29, 441-450.

Taguchi, G. (1987), *System of Experimental Design*, White Plains: Unipub/Kraus International Publications.

Acknowledgments

Appendix A: The Science of Bubbles

Bubbles or soap films are no trivial matter. For example, Boys (1959) and Isenberg (1992) give elementary accounts of the science that is involved. What follows is a brief glimpse of bubble solutions at a molecular level.

Water is extraordinarily sticky because of hydrogen bonding in which the molecules of water are electrically attracted to each other. The well-known H_2O formula for water indicates that a molecule of water consists of two hydrogen atoms and one oxygen atom; the two hydrogen atoms are situated on one side of the molecule and are attracted to the oxygen atoms of other nearby water molecules. Moreover, the electrical attraction can be explained at an atomic level; the single hydrogen electrons spend more time inside the water molecule toward the oxygen atom that makes the hydrogen side positively charged. Also, the eight oxygen electrons spend more time away from the two hydrogen atoms, which makes the oxygen side negatively charged. Thus, hydrogen bonding can be explained because opposite charges attract. This attraction is needed because the constant thermal agitation of the water molecules would otherwise cause the water to boil instantly or evaporate. The thermal agitation also prevents the water molecules from being locked into a fixed pattern, but the numerous temporary hydrogen bonds between the water molecules make water an extraordinarily sticky fluid. Surface tension plays an important role in the way in which a liquid behaves. Within water at more than a few molecules away from its surface, molecules are in equilibrium in that they feel no net force from their surrounding neighbors. At the surface, however, the molecules have no neighbors pulling from above so that the surface molecules tend to be pulled back into the liquid. When a bubble is blown, the surface is stretched and increases in area as more molecules are dragged from within the liquid to the surface; this stretching effect is known as surface tension.

Trying to blow a bubble with plain water will not work. There is a common misconception that water does not have the necessary surface tension to maintain a bubble. In fact, the opposite is true. The surface tension in plain water is just too strong for bubbles to last for any length of time. Moreover, water evaporates so that the surface quickly becomes thin, causing them to pop. Now consider soap molecules that consist of long chains of carbon and hydrogen atoms. The atoms are configured so that one end of the chain likes to be in water (or hydrophilic) while the other is greasy (or hydrophobic) and wants to avoid water. In a soap-and-water solution the greasy ends of the soap molecule find their way to the surface and squeeze their way between the surface water molecules. With the greasy end sticking out of the water, the surface water molecules are separated from each other, thereby, reducing the surface tension force as much as a third of plain water. Because the greasy ends of the soap molecules stick out of the bubble surface, the soap film is protected somewhat from evaporation (since grease is less prone to evaporation) and thus substantially prolongs the life of the bubble.

Appendix B: The Bubble Mixture Experiment Data

Mixture Number	Run Order	Proportions Water	Proportions Soap	Proportions Glycerin	Soap Type	Water Type	1	2	3	4	5	Total
1	2	0.6	0.35	0.05	-1	-1	19	25	25	25	25	119
2	3	0.6	0.25	0.15	-1	-1	13	10	14	12	14	63
3	20	0.65	0.35	0	-1	-1	9	9	7	6	8	39
4	32	0.65	0.25	0.1	-1	-1	16	7	12	13	12	60
5	1	0.7	0.25	0.05	-1	-1	11	13	15	9	16	64
6	15	0.7	0.15	0.15	-1	-1	7	11	11	13	12	54
7	41	0.775	0.145	0.08	-1	-1	16	12	10	4	2	44
8	24	0.8	0.2	0	-1	-1	11	8	13	9	10	51
9	31	0.81	0.04	0.15	-1	-1	1	2	2	5	2	12
10	29	0.85	0.12	0.03	-1	-1	17	13	9	14	11	64
11	23	0.88	0.04	0.08	-1	-1	5	1	2	1	1	10
12	27	0.95	0.05	0	-1	-1	3	5	6	2	4	20
1	34	0.6	0.35	0.05	-1	1	13	15	13	13	17	71
2	40	0.6	0.25	0.15	-1	1	15	15	15	13	10	68
3	7	0.65	0.35	0	-1	1	9	14	11	14	5	53
4	17	0.65	0.25	0.1	-1	1	16	15	10	14	5	60
5	6	0.7	0.25	0.05	-1	1	16	15	18	17	11	77
6	35	0.7	0.15	0.15	-1	1	14	7	6	10	9	46
7	36	0.775	0.145	0.08	-1	1	7	11	11	13	14	56
8	48	0.8	0.2	0	-1	1	14	12	16	12	14	68
9	8	0.81	0.04	0.15	-1	1	4	4	4	4	4	20
10	18	0.85	0.12	0.03	-1	1	4	6	10	11	9	40
11	4	0.88	0.04	0.08	-1	1	1	9	2	6	5	23
12	37	0.95	0.05	0	-1	1	2	3	2	4	3	14
1	43	0.6	0.25	0.15	1	-1	8	2	4	3	4	21
2	44	0.6	0.35	0.05	1	-1	6	2	1	3	1	13
3	10	0.65	0.25	0.1	1	-1	3	2	2	3	2	12
4	26	0.65	0.35	0	1	-1	3	6	8	6	6	29
5	5	0.7	0.25	0.05	1	-1	1	4	4	10	1	20
6	42	0.7	0.15	0.15	1	-1	3	2	4	6	1	16
7	47	0.775	0.145	0.08	1	-1	2	7	8	3	2	22
8	38	0.8	0.2	0	1	-1	1	6	5	3	4	19
9	33	0.81	0.04	0.15	1	-1	4	5	1	1	4	15
10	13	0.85	0.12	0.03	1	-1	3	3	0	1	1	8
11	19	0.88	0.04	0.08	1	-1	0	1	0	1	1	3
12	25	0.95	0.05	0	1	-1	1	2	3	6	3	15
1	11	0.6	0.25	0.15	1	1	5	3	5	2	3	18
2	46	0.6	0.35	0.05	1	1	5	4	3	0	5	17
3	39	0.65	0.25	0.1	1	1	0	1	1	2	1	5
4	45	0.65	0.35	0	1	1	2	5	2	1	3	13
5	12	0.7	0.15	0.15	1	1	3	7	2	5	5	22
6	22	0.7	0.25	0.05	1	1	1	2	2	6	6	17
7	14	0.775	0.145	0.08	1	1	8	4	1	1	1	15
8	30	0.8	0.2	0	1	1	3	4	1	3	0	11
9	21	0.81	0.04	0.15	1	1	2	5	6	5	3	21
10	16	0.85	0.12	0.03	1	1	2	4	4	3	4	17
11	9	0.88	0.04	0.08	1	1	0	1	0	0	2	3
12	28	0.95	0.05	0	1	1	4	2	5	3	1	15

For soap type: –1 = Joy and +1 = Ivory
For water type: –1 = spring and +1 = tap

Appendix C: The R Output for the Reduced Quadratic Model

glm(formula = y1 ~ -1 + W + S + G + WS + WG + SG + Wn1 + Sn1 + Gn1 + WSn1 + WGn1 + SGn1, family = poisson)   %n1 corresponds to N_s

Model: poisson, link: log

	Df	Deviance	Resid. Df	Resid. Dev
NULL	48	8701.7		
W	1	7033.0	47	1668.7
S	1	909.1	46	759.5
G	1	79.2	45	680.4
WS	1	24.9	44	655.5
WG	1	7.8	43	647.6
SG	1	7.5	42	640.1
Wn1	1	433.3	41	206.8
Sn1	1	52.5	40	154.3
Gn1	1	0.3	39	154.0
WSn1	1	1.3	38	152.7
WGn1	1	1.6	37	151.0
SGn1	1	31.0	36	120.0

Deviance Residuals:

Min	1Q	Median	3Q	Max
-2.7643	-1.0658	-0.1526	0.8362	4.6545

Coefficients:

	Estimate	Std.	Error	z-value	Pr(>|z|)	
W	2.3938	0.1703	14.054	<2e-16	***
S	-3.3553	2.1339	-1.572	0.1159	
G	17.7969	10.3578	1.718	0.0858	
WS	13.9705	3.5458	3.940	8.15e-05	***
WG	-17.5638	12.3313	-1.424	0.1544	
SG	-5.2577	12.5849	-0.418	0.6761	
Wn1	-0.1587	0.1703	-0.932	0.3515	
Sn1	2.2537	2.1339	1.056	0.2909	
Gn1	47.2846	10.3578	4.565	4.99e-06	***
WSn1	-5.2169	3.5458	-1.471	0.1412	
WGn1	-53.1718	12.3313	-4.312	1.62e-05	***
SGn1	-69.7290	12.5849	-5.541	3.01e-08	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Dispersion parameter for poisson family taken to be 1)

Null deviance: 8701.69 on 48 degrees of freedom
Residual deviance: 120.03 on 36 degrees of freedom
AIC: 385.12

Number of Fisher Scoring iterations: 4