Applications of Experimental Design in Engineering and Food Science

In Partial Fulfillment of the Requirements for the Degree

Bachelor of Science, Statistics

Hector R Herrera

December 2013

Contents

Part I Application in Engineering:

Introduction:

In collaboration with an engineering firm in Santa Barbara County, the department of Statistics at California Polytechnic sought to establish aesthetic performance tracking, control, and capability of a new product line. The new product line required a finish painting operation with stringent customer requirements for coating and aesthetic performance. Through data collection and monitoring, paint operation parameters would be optimized in order to reduce the amount of production parts that do not meet guidelines along the production line. Process optimization should ultimately lead to decreased production costs by reducing the operation's cost of quality. The proposed team would be comprised of one statistics student, five materials engineers and two Cal Poly faculty members; one from each department. While this project is ongoing, a portion of the project was completed and is reported in this document. Additionally, participants of this collaboration are under a nondisclosure agreement and proprietary information may not be revealed or released to anyone not directly involved in the project.

Background:

The proposed project is broken into six sub projects which include measurement instrumentation system analysis and process quality control (gage study and paint design of experiment), witness panel sampling, film thickness determination, paint process mapping and inspection gate establishment, paint process characterization and finally paint process monitoring and control. Results from the initial stages of this larger project are reported here, including measurement instrumentation system analysis and process quality control, including a gage study and paint process design of experiment to characterize sources of variability for paint thickness and aesthetic performance.

Gage Study:

The painting process is characterized by measurements of the final painted surfaces. Automated measurements of film thickness, gloss and color are made using several instruments. In order to rely on any results from our study, the instruments used to make measurements must be tested to ensure they are capable of first, measuring at the specification tolerances that the production parts are held to, and second, able to reproduce the measurement. To achieve this task we performed a Gage R&R Study (repeatability and reproducibility). Repeatability refers to the variation in measurements taken by a single person or instrument on the same item under the same conditions while reproducibility refers to the variation induced when different operators, instruments, or laboratories measure the same item.

In this study we looked at two instruments, across three appraisers, 24 parts, 3 measurements per part and a total of 20 different measurements were analyzed. These measurements include paint thickness, effects (sparkle- intensity, grade, area and diffusion), and color. According to our study all 20 aspects of the devices were within working range (calculated as a gage to part ratio- GPR) and did not induce more variability than expected.

Since no specific guidelines were in place to measure the gage's allowable variance, we used the gage to part ratio as a measure of the percentage of a part's variability that can be attributed to the gage itself. This measure was calculated as follows:

$$\frac{GRR}{\overline{XD}} \times 100\%$$

GRR refers to the Repeatability and Reproducibility which is calculated as follows:

$$\sqrt{Equipment\ Variation^2 + Appraiser\ Variation^2}$$

$\overline{XD}$ refers to the mean measurement across three appraisers for all 24 samples.

Equipment variation (EV-Repeatability) is calculated as follows:

$$\overline{RD} \times K_1$$

Where $\overline{RD}$ is the mean range of the measurement being checked, across all 24 samples and 3 operators. K_1 is the established gage study constant for the number of trials of each part conducted (three in this experiment).

Appraiser variation (AV-Reproducibility) is calculated as follows:

$$\sqrt{\left(\bar{X}_{Diff} \times K_2\right)^2 - \frac{EV^2}{n \times r}}$$

$\overline{X}_{Diff}$ is the maximum difference between the mean measurements of the 24 parts being tested. K_2 is the established gage study constant for the number of operators (three in this experiment). n the number of samples (24), and r the number of trials (3).

Figure 1 is a sample of the Gage R&R for the thickness measurement looking at the interaction of part number and operator. While Figure 2 illustrates a plot of the thickness measurement by operator. In both cases we notice no major differences between part number and operators, indicative of a functioning gage.

Figure 1:

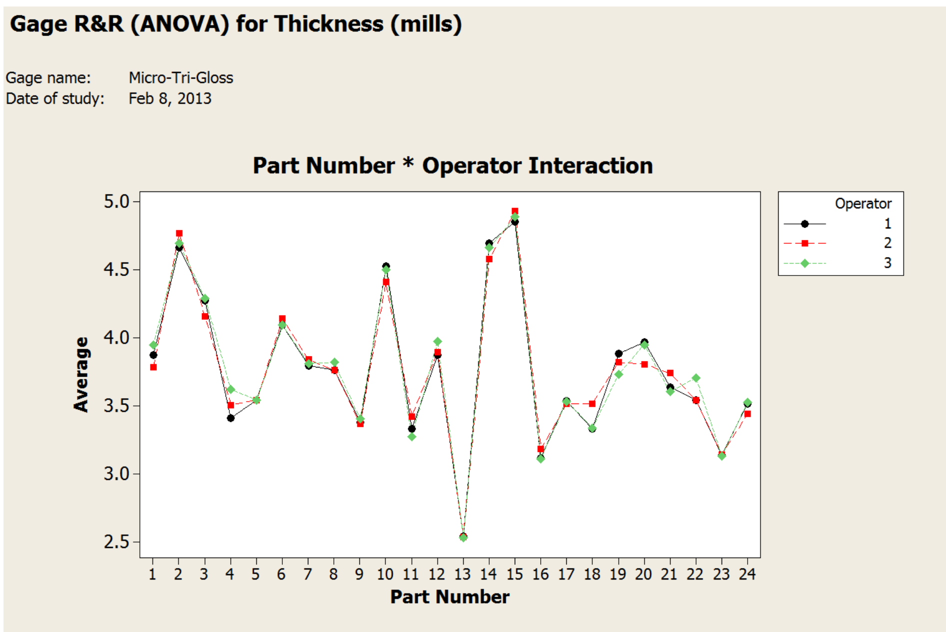
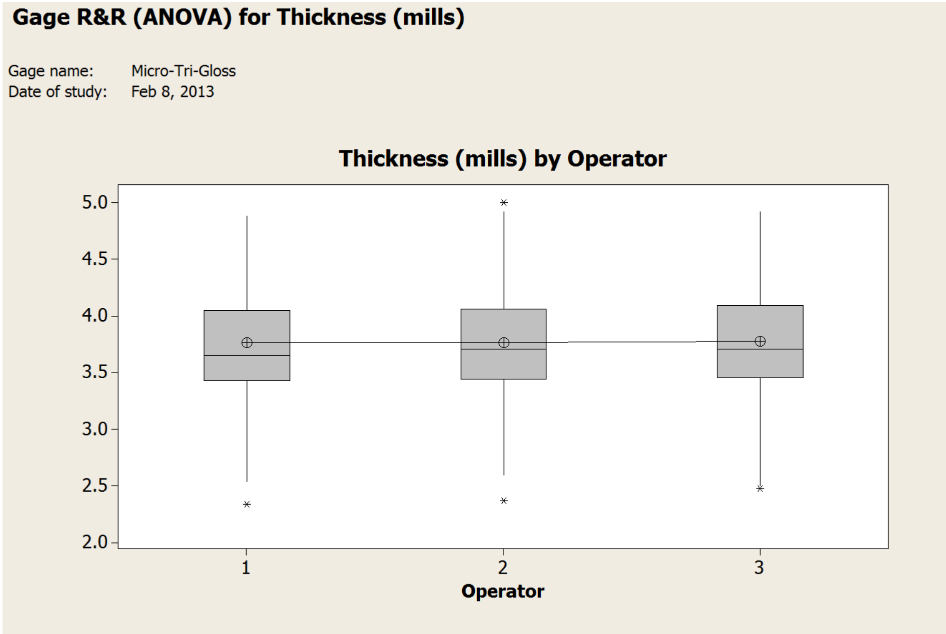


Figure 2:



Conclusion:

While the project is ongoing at the production facility, we were able to successfully complete a gage study and a design of experiment. We conclude that both the measurement instruments and the physical paint process are able to measure and produce parts that are within specifications. Future work with the company could be achieved by picking up the project where we left off and start measurements of production parts to establish and implement part guidelines.

Part II Applications in Food Science:

Introduction:

In collaboration with the Department of Food Science at Cal Poly, the team is comprised of a graduate student working on her thesis, a statistics student working on his senior project (author of this document), and their respective advisers, seek out to answer the following three research questions through a food sensory panel conducted on the Cal Poly campus:

1. Is there a relationship between a person's BAS score and salt/fat levels in cheeses?
2. What product attributes affect the end emotional state of a panelist?
3. Is product liking better modeled when emotion is taken into account compared to hedonics alone?

Background:

Using a panelist's emotion information is a novel method for capturing overall product acceptance or preference in sensory. This study looks at 8 convenience string cheeses in the form of mozzarella, mixed mozzarella, and particulate mozzarella cheeses. The study is comprised of 77 members of the Cal Poly community, which include students, faculty, and city of San Luis Obispo residents.

In traditional sensory panels a participant is asked to sample a product, then using computer aided software rate the product on a range of attributes. These attributes include: overall liking, flavor liking, texture liking, appearance liking, aroma liking, and aftertaste liking; all measured on a one to nine discrete scale where one is *extremely dislike* and nine is *extremely like*. The six previously mentioned attributes will be referred to as product *hedonics* in the remainder of this paper. Additionally there are 16 product texture attributes a panelist is asked to describe for each product. The textures attributes scale range from zero to five and vary by the specific texture attribute being asked, texture includes both quantitative and categorical measurements. These texture attributes include: bite location, break resistance, bounciness, fibrosity, shear, resilience, stickiness, surface deviation, grittiness, crumble, soft, sponginess, rubbery, sandy, hardness, and waxy.

A panelist's BAS score refers to a questionnaire that looks at personality dimension of behavioral responses to appetitive cues and is known as the behavioral approach system, or BAS for short. This questionnaire is comprised of 24 questions measured on a Likert scale ranging from one to four, where a 1 is *very true* and a 4 is *very false*. To calculate an individual's BAS score a simple sum of the responses is used. Previous research has shown that BAS scores are positively correlated with emotional overeating, binging, preference for fatty and sweet foods, neural responses to food pictures, and with general dysfunctional eating.

In addition to traditional hedonic and texture attributes, this study looks at the emotional state of a panelist during the testing session and attempts to quantify the emotional state of the panelist. Data is collected on seven individual emotions which include: excited, sociable, self-confident, fatigued, judgmental, raging, and sad. For each emotion and before each sample a panelist is asked to quantify their current emotional level. Additionally, after consumption of each product, a panelist is again asked to quantify their emotional state. This scale ranges from zero to five where zero is *not at all* and five is *extremely*.

When assessing products, the measure of overall liking is used as an indicator of general product acceptance and as an indicator of whether a consumer might purchase/consume the product in question again. As a result, a higher measure of overall end liking is assumed to be a better liked product compared to others.

Study Design:

Stage one of the study includes recruitment as well as some pre-sensory day work from the participants. Each participant is asked to complete a 7 x 5 grid poster board, one column for each of the seven emotions, each with five levels ranging from one to five. In each of the grid spots, the panelist is asked to find a picture that best represents that emotional state, to use as a reference during the sensory test. Additionally each panelist is asked to complete and turn in the BAS questionnaire.

On the day of testing, each panelist will consume a total of eight commercial convenience string cheeses. The order of consumption is randomized for each panelist, except for cheese sample number 8, since it was a flavored, particulate cheese. Due to not being randomized the last cheese sample is not included in this analysis.

Before each sample, panelists are asked to score their emotional state. They then proceed to take an initial bite of the sample and answer hedonic and texture questions. At the end of each sample they are asked again the hedonic, texture, and emotional state, resulting in two separate scores for hedonics, texture, and emotion for each of the eight samples. The process is repeated for all eight samples with a palate cleanser between samples.

Analysis & Results:

Paired t-tests are used to assess significance of the two hedonic (and texture) scores for each sample. Since there are no significant changes between the two a simple average is calculated for each hedonic and texture attributes, for continuous variables. Due to coding errors and unclear directions during testing several texture attributes are excluded from the analysis, reducing total texture attributes to 13 analyzed attributes. A total of 78 panelists participated. The final analysis includes 76 participants when addressing the BAS scale response research question and 77 for all remaining analysis. The reason for the disagreement is due to a non-completed BAS questionnaire by one panelist. One panelist was removed from all analyses due to incomplete responses.

Research Question 1: Is there a relationship between a person's BAS score and salt/fat levels in cheeses?

The study design is not a full factorial and as a consequence the interaction between salt and fat cannot be addressed; the breakdown of observations by salt and fat content can be seen in Table 1. An analysis of covariance (ANCOVA) was conducted to determine the relationship between overall product liking, BAS scores, and salt/fat content. The five assumptions of ANCOVA were satisfied for these analyses. Salt and fat contents are coded as categorical, BAS score is continuous, and panelist is treated as a random effect. The following models are used:

(1.1)

$$Overall\ Liking = \beta_0 + \beta_1 BAS + \beta_2 Salt + \beta_3 Panelist$$

(1.2)

$$Overall\ Liking = \beta_0 + \beta_1 BAS + \beta_2 Fat + \beta_3 Panelist$$

Table 1: Breakdown of Number of Observations by Fat/Salt Content

	% Fat					
% Salt	10.4	16.1	19	20.8	21.4	33.3
0.625						76
0.677			76			
0.679		76			76	
0.708				152		
0.833	76					

Table 2: Least Squares Means and Letter's Plot for Salt Content

Grouping				
% Salt				Mean
0.708	A			6.60
0.833	A			6.18
0.625	A			5.93
0.679		B		5.01
0.677			C	4.17

Table 3: Least Squares Means and Letter's Plot for Fat Content

Grouping				
% Fat				Mean
20.8	A			6.60
10.4	A			6.18
33.3	A			5.93
21.4		B		5.01
16.1		B		5.00
19			C	4.17

Results:

In both models 1.1 and 1.2 statistical significance is achieved for the predictors salt and fat, p-values <.0001 in both cases. The p-value for the BAS score in both models is .1023, inconclusive evidence as to whether BAS score influences product liking. While BAS scores does appear to give weak evidence of overall product liking, this result could be attributed to a lack of power and further testing may be considered. Furthermore based on Tukey's letter plots we notice we do not have a linear relationship between salt and fat contents, displayed in Tables 2 & 3. We are able to detect differences between the salt and fat levels, but these differences could be attributed to the type of cheese being sampled.

Research Question 2: What product attributes affect the end emotional state of a panelist?

To address this question a series of MANOVA models are considered. Since it is very likely that a panelist's current emotional state depends on their other emotions, i.e. it is not likely that you will be both highly raging and highly sociable, MANOVA will best model these interrelated emotional responses. For the purpose of this analysis we consider three models, in all models we account for the product and panelist. The first model looks at emotions alone, using the initial emotional state to predict the end emotional state (2.1). The second model looks at the hedonic attributes that affect end emotional state (2.2). And the third model addresses the question of texture attributes and how they affect emotion (2.3). A final model is also checked that includes both hedonic and texture attributes in predicting end emotion (2.7). For this set of models and the remaining analyses, panelists are treated as fixed factors (due to the recommendation of the research leader and an ongoing debate in the food science world). A point to notice is that the negative emotions are skewed right; to address this question a log transformation was considered but resulted in more extreme (smaller) p-values, so the non-transformed emotion values are used.

(2.1)

$$
\begin{aligned}
&End\ Emotion\ (7\ levels) \\
&\qquad = \beta_0 + \beta_{1\ldots7} Initial\ emotion\ (7\ levels) + \beta_8 Sample\ Name \\
&\qquad + \beta_9 Panelist
\end{aligned}
$$

(2.2)

$$End\ Emotion\ (7\ levels) = \beta_0 + \beta_{1\ldots5}Hedonics + \beta_6 Sample\ Name + \beta_7 Panelist$$

(2.3)

$$End\ Emotion\ (7\ levels) = \beta_0 + \beta_{1\ldots13}Texture + \beta_{14} Sample\ Name + \beta_{15} Panelist$$

In addition to these three models, an additional three models (2.4-2.7) are fit that include principal components for hedonic and texture attributes. Two principal components are selected for both initial and end emotional state and account for 59% and 60% of the variation in emotional responses, Table 4. Two principal components are selected from the five hedonic attributes and account for 84% of the variability in hedonic responses, Table 5. For texture attributes four principal components are selected from the 13 attributes and account for 60% of the variability in texture responses, Table 6.

(2.4)

$$End\ Emotion\ PC\ (2\ levels) = \beta_0 + \beta_{1,2}Emotion\ PCs + \beta_3 Sample\ Name + \beta_4 Panelist$$

(2.5)

$$End\ Emotion\ (7\ levels) = \beta_0 + \beta_{1,2}Hedonic\ PCs + \beta_3 Sample\ Name + \beta_4 Panelist$$

(2.6)

$$End\ Emotion\ (7\ levels) = \beta_0 + \beta_{1\ldots4}Texture\ PCs + \beta_5 Sample\ Name + \beta_6 Panelist$$

(2.7)

$$End\ Emotion\ (7\ levels) = \beta_0 + \beta_{1,2}Hedonic\ PCs + \beta_{3,4}Texture\ PCs + \beta_5 Sample\ Name + \beta_6 Panelist$$

Table 4: Loading Matrix for Principal Components (Emotions)

Emotion- 61.5%	Prin1	Prin2
Excited	0.74	-0.48
Sociable	0.73	-0.52
Self-Confident	0.65	-0.46
Fatigued	0.36	0.45
Judgmental	0.43	0.48
Raging	0.61	0.54
Sad	0.52	0.65

Table 5: Loading Matrix for Principal Components (Hedonics)

Hedonics-84%	Prin1	Prin2
Flavor Liking	0.85	-0.34
Texture Liking	0.84	-0.30
Appearance Liking	0.78	0.59
Aroma Liking	0.80	-0.22
Aftertaste Liking	0.92	0.28

Table 6: Loading Matrix for Principal Components (Texture)

Texture-60%	Prin1	Prin2	Prin3	Prin4
Break Resistance	0.44	-0.55	0.14	0.13
Fibrousity	0.52	-0.14	0.05	0.52
Shear	-0.14	0.07	0.66	0.58
Stickiness	0.37	0.51	0.22	-0.22
Surface Deviation	0.50	0.11	-0.24	-0.06
Grittiness	0.73	0.21	-0.21	0.23
Crumble	0.70	0.30	-0.10	0.07
Soft	-0.30	0.71	0.41	-0.06
Sponginess	0.48	0.35	0.42	-0.23
Rubbery	0.44	-0.53	0.45	-0.18
Sandy	0.66	0.26	-0.22	0.02
Waxy	0.45	-0.39	0.29	-0.45

Results for Individual Predictors:

MANOVA p-values presented are Wilk's Lambda. Model 2.1 results in a MANOVA full model p-value of .037. To address how initial emotions affect end emotions, contrasts are constructed using only the paired initial and end emotions. Using the contrasts we find that the initial emotions of *Excited* and *Sad* are significant predictors of end emotional state after we account for other emotions, contrast p-values of .01 and .03 respectively. Furthermore emotion *Judgmental* is almost significant at the .05 level with a p-value of .053.

Model 2.2, which addresses hedonic attributes, has a whole model p-value of <.01, results in a borderline significant predictor *flavor liking* (p-value of .054). Each emotion was also separately checked using an ANOVA model and we find that flavor is a significant predictor in *Excited* (p-value <.01), *Judgmental* (p-value = .02) and *Sad* (p-value =.05) in the case of the positive emotion we find that an increase in flavor score results in an increase in *Excited* and an increase in flavor score results in a decrease in negative *Judgmental* and *Sad* emotions. Texture was found to be a significant predictor of *Excited* (p-value =.01), *Sociable* (p-value =.04), *Self-Confident* (p-value =.05), *Judgmental* (p-value =.05), and *Sad* (p-value =.04). When addressing the positive emotions we notice that higher scores on texture liking result in higher positive emotions and decrease in negative emotions. The final two hedonic attributes that are significant predictors of emotions are, aftertaste, which affects *Sociable* (p-value =.02) in a positive direction, and appearance, which affects *Judgmental* (p-value =.05) in a positive direction as well.

Whole model p-value for model 2.3 is <.01 with predictors sponginess (p-value .05), rubbery (p-value <.01), and waxy (p-value .04) are found to be significant in the MANOVA. Additionally each emotion was tested separately and several texture attributes were found to be predictors of emotions, with waxy appearing in five of the seven emotions. For purposes of brevity the full results of texture will not be discussed in this paper.

Results for Principal Component Predictors:

The whole model MANOVA p-value for model 2.4 is .08, indicating little evidence to suggest our principal components are adequate indicators of end emotional state. No additional analysis was considered for this specific model.

In model 2.5 we address the seven end emotions using two principal components for the hedonic attributes. Full MANOVA p-value of <.01 indicate that both components are significant predictors of emotion (p-values <.01). Separately we analyze each emotion and the resulting p-values can be seen in Table 7.

Table 7: P-Values for Individual Principal Components on Emotions

	Excited	Sociable	Self-Confident	Fatigued	Judgmental	Raging	Sad
PC1	<0.01	<0.01	0.01	0.01	<0.01	<0.01	<0.01
PC2	<0.01	0.01	0.10	0.05	<0.01	0.35	0.01

Using texture principal components the whole model MANOVA (model 2.6) p-value <.01 indicates the texture principal components model emotions well. We do not, however, observe significant predictors in the whole model. The principal component four with p-value of .10 is the closest to achieving significance at the .05 level.

In the final MANOVA model, 2.7, with whole model p-value of <.01 where we address both, principal components of hedonics and texture on emotion, we find that both principal components of hedonics are significant (p-values both <.01), but not any principal components of texture. It is worth noting that principal component 4 of the texture attributes is almost significant, with p-value .08. From this study we conclude that hedonics are much more likely to drive changes in emotion than are texture attributes of food.

Research Question 3: Is product liking better modeled when emotion is taken into account compared to hedonics alone?

To address the third research question two models are constructed (3.1 and 3.2). To avoid a cumbersome model we will build models using the principal components of both hedonics and emotion only, rather than 17 individual predictors. In these models the response variable is end overall product liking. We are using the measure as an indicator of product acceptance. Model 3.1 looks at a model with only the principal components of Hedonics while model 3.2 looks at both principal components of hedonics as well as those relating to initial emotion. ANOVA assumptions are met in these set of models.

(3.1)

$$End\ Overall\ Liking = \beta_0 + \beta_{1,2} Hedonic\ PCs + \beta_3 Panelist + \beta_4 Sample$$

(3.2)

$$\begin{aligned} End\ Overall\ Liking \\ &= \beta_0 + \beta_{1,2} Hedonic\ PCs + \beta_{3,4} Emotion\ PCs + \beta_5 Panelist \\ &+ \beta_6 Sample \end{aligned}$$

From Table 6 we notice that the adjusted R^2 between the two models is roughly the same. When we compare the model sum of squares the difference is negligent. We can conclude that emotions do not provide additional insight into product liking versus hedonics alone.

Table 6: Models 3.1 and 3.2 Model Comparisons with Effect P-Values

	Model 3.1	Model 3.2
Adjusted R^2	0.86	0.86
SS Model	2106.37	2107.9
SS Error	278.36	276.82
Model P-Values		
Panelist	<0.01*	<0.01*
Sample	0.08	0.07
PC1 (hedonics)	<0.01*	<0.01*
PC2 (hedonics)	<0.01*	<0.01*
PC1 (emotion initial)	-	0.68
PC2 (emotion initial)	-	0.12

Conclusion:

In summary, each of the three research questions and conclusions are restated below.

1. Is there a relationship between a person's BAS score and salt/fat levels in cheeses?

   Based on this study, a person's BAS score is not a significant predictor of overall product liking when we account for the salt and fat levels in cheeses. We are able to detect differences between the salt and fat levels, but these differences could be attributed to the type of cheese being sampled.

2. What product attributes affect the end emotional state of a panelist?

   Overall when we address the question of hedonics and texture on emotional responses, we find that hedonics alone are better predictors of changes in emotion when including textural attributes; seen in model 2.7. Looking at the specific hedonics that affect emotional state, model 2.2, we find that flavor is a significant predictor in *Excited*, *Judgmental* and *Sad*. For the positive emotion, *Excited*, we find that an increase in flavor score results in an increase in the end emotion *Excited.* An increase in flavor score results in a decrease in end *Judgmental* and end *Sad* emotions. Texture was found to be a significant predictor of *Excited*, *Sociable*, *Self-Confident*, *Judgmental*, and *Sad*. When talking about the positive emotions we notice that higher scores on texture liking results in higher end positive emotions and decrease in end negative emotions. The final two hedonic attributes that were significant predictors of emotions are aftertaste that affect *Sociable* in a positive direction and appearance which affects *Judgmental* in a positive direction as well.

3. Is product liking better modeled when emotion is taken into account compared to hedonics alone?

   The results of this study do not provide evidence for the hypothesis that emotions would better represent overall product liking and acceptance. Again, we should take note that it is not possible that emotions are providing information towards product liking, only that we do not have evidence in this study to support the theory. Additional testing will be necessary to provide evidence for this theory.

Future Directions:

There is a study conducted by the Department of Food Science at Cal Poly in which salt substitutes in cheeses are used. Additionally, there is a completed data set which has an added calibration step for emotion. The idea is that, traditionally, after each sample we cleanse the palate and thus should attempt to cleanse the emotional ‘palate’ as well. Participants for that study were again asked to complete a BAS questionnaire, a poster board with the seven scaled emotions and in addition were asked to provide an additional image, one that is calming and is joyful to look at. Analysis on this data set will commence shortly.

Currently an additional study is in the design stage where face-reading software is being implemented. For the current studies we have been asking participants to make a cognitive assessment on a subconscious response, with the face-reading software we will be able to ensure that the response participants are providing actually match what they respond on questionnaires.