This analysis report is concerned with the relationship between working conditions and smoking on biological levels of cotinine. The variables “priming”, “barning”, “topping”, and “other” are indicator variables describing work tasks. The test of whether all cell means for task are equal in the model

$$\text{log salivary cotinine} = \beta_0 \bullet \text{priming} + \beta_1 \bullet \text{barning} + \beta_2 \bullet \text{topping} + \beta_3 \bullet \text{other}$$

is equivalent to testing H_0: $\beta_0=\beta_1=\beta_2=\beta_3$. We failed to reject the null (F-stat = 116.20, df = 3, p-value < 0.0001) and conclude the cell means for the tasks are not equal. Pairwise comparisons were done using Scheffe’s method, the results of which are tabulated below.

Least Squares Means for Effect task

i	*j*	*Difference Between Means*	*Simultaneous 95% Confidence Limits for LSMean(i)-LSMean(j)*		*Degrees of freedom*	*F-statistic*	*p-value*
1	*2*	0.920781	0.337572	1.503991	1, 690	19.57599	0.0002
1	*3*	1.673848	1.265368	2.082328	1, 690	131.8714	< 0.0001
1	*4*	2.699252	2.284530	3.113975	1, 690	332.6837	< 0.0001
2	*3*	0.753067	0.167517	1.338616	1, 690	12.98968	0.0049
2	*4*	1.778471	1.188549	2.368393	1, 690	71.37796	< 0.0001
3	*4*	1.025404	0.607397	1.443411	1, 690	47.25875	< 0.0001

Table 1

As can be seen in Table 1, the differences in mean values between each group are all highly significant, where group numbers correspond to the order of tasks listed in the model above.

Parameter	*Estimate*	*Standard Error*
priming	4.508556597	0.10222010
barning	3.587775112	0.18127655
topping	2.834708541	0.10390980
other	1.809304310	0.10701227

Table 2

Parameter	*Estimate*	*Standard Error*
Intercept	4.508556597	0.10222010
barning	-0.920781485	0.20811087
topping	-1.673848056	0.14576075
other	-2.699252287	0.14798843

Table 3

Tables 2 and 3 display the parameter estimates using cell mean coding and reference cell coding respectively, with task priming as the reference category.

If we wished to test the hypothesis that the average cotinine level for priming workers exceeds that of all other workers, the contrast matrix used would have to account for the different sizes of the barning (n = 69), topping (n = 210) and other (n = 198) task groups so works out to be C = [1 -0.145 -0.440 -0.415] with θ_0=0 under cell mean coding, while under reference cell coding C = [0 -0.145 -0.440 -0.415] with the same θ_0.

A two-way ANOVA model was fit using a reference cell coding scheme, adding a binary variable “wet” to the previous model, including interaction terms:

$$\text{log salivary cotinine} = \beta_0 + \beta_1 \bullet \text{barning} + \beta_2 \bullet \text{topping} + \beta_3 \bullet \text{other} + \beta_4 \bullet \text{wet} + \beta_5 \bullet \text{barning} \bullet \text{wet} + \beta_6 \bullet \text{topping} \bullet \text{wet} + \beta_7 \bullet \text{other} \bullet \text{wet}$$

The interpretation of the estimate of β_0 is the mean log salivary cotinine level of a priming worker in dry conditions. The estimate of β_1 is the change in mean log salivary cotinine level comparing a barning worker in dry conditions to a priming worker in dry conditions. Similarly, the estimates of β_2 and β_3 are analogous for topping and other respectively. The estimate of β_4 is the change in mean log salivary cotinine level comparing a priming worker in wet conditions to one in dry conditions. Similarly, the estimates of β_5, β_6 and β_7 are analogous for barning, topping and other respectively.

The validity of the HILE Gauss assumptions were confirmed primarily by residual analysis. The histogram and QQ plot confirmed the Gaussian assumption and a scatterplot confirmed homoscedacity. Independence was assumed by study design, existence is guaranteed by finite sample and linearity is implicit in ANOVA.

Table 4 shows the mean values for each of the eight combinations of task and wet.

Parameter	*Estimate*	*Standard Error*	*Coefficients*
Marg Mean: Priming Dry	4.26933727	0.18547121	β_0
Marg Mean: Barning Dry	3.54274757	0	$\beta_0 + \beta_1$
Marg Mean: Topping Dry	2.68818516	0	$\beta_0 + \beta_2$
Marg Mean: Other Dry	1.80888875	0	$\beta_0 + \beta_3$
Marg Mean: Priming Wet	4.61311604	0	$\beta_0 + \beta_4$
Marg Mean: Barning Wet	3.66702358	0	$\beta_0 + \beta_1 + \beta_4$
Marg Mean: Topping Wet	2.87930261	0	$\beta_0 + \beta_2 + \beta_4$
Marg Mean: Other Wet	1.81053437	0	$\beta_0 + \beta_3 + \beta_4$

Table 4

Another model was fit using a reference cell coding scheme, adding a continuous variable "lnnsmoke" (defined as log[1 + # of cigarettes smoked]) to the previous model, including interaction terms:

$$\text{log salivary cotinine} = \beta_0 + \beta_1 \bullet \text{barning} + \beta_2 \bullet \text{topping} + \beta_3 \bullet \text{other} + \beta_4 \bullet \text{wet} + \beta_5 \bullet \text{lnnsmoke} + \beta_6 \bullet \text{barning} \bullet \text{wet} + \beta_7 \bullet \text{topping} \bullet \text{wet} + \beta_8 \bullet \text{other} \bullet \text{wet} + \beta_9 \bullet \text{barning} \bullet \text{smoke} + \beta_{10} \bullet \text{topping} \bullet \text{smoke} + \beta_{11} \bullet \text{other} \bullet \text{smoke}\ \beta_{12} \bullet \text{wet} \bullet \text{smoke} + \beta_{13} \bullet \text{barning} \bullet \text{wet} \bullet \text{smoke} + \beta_{14} \bullet \text{topping} \bullet \text{wet} \bullet \text{smoke} + \beta_{15} \bullet \text{other} \bullet \text{wet} \bullet \text{smoke}$$

The 16 parameter estimates are listed in Table 5 below

Parameter	*Estimate*	*Parameter*	*Estimate*
Intercept	3.883414621	*other_wet*	-0.327038865
barning	-0.888324180	*barning_smoke*	0.287856212
topping	-2.066780957	*topping_smoke*	0.823713934
other	-3.050019611	*other_smoke*	1.335608893
wet	0.611992088	*wet_smoke*	-0.283108376
lnnsmoke	0.510018497	*barning_wet_smoke*	0.024005208
barning_wet	-0.266139213	*topping_wet_smoke*	0.148209637
topping_wet	-0.355995374	*other_wet_smoke*	0.050316835

Table 5

Rather than interpret each parameter, a representative one is selected, with other interpretations somewhat analogous given the full model listed above. The estimate of β_{14} is the expected increase in log salivary cotinine with one unit increase in log(1 + # cigarettes smoked) for a topping worker in wet conditions *beyond the contributions* due to smoking, the interaction of smoking and topping, and the interaction between smoking and wet.

For the ANCOVA model to hold, each interaction term with lnnsmoke would have to be insignificant but the interactions with task are highly significant (example is estimate of β_{11}, t-stat = 6.89, p-value < 0.0001). Therefore the ANCOVA model does not hold.

Group-wise backwards selection was chosen as the method to optimize the model. The first group tested were the 3-way interactions, followed by 2-way interactions involved "wet". The selected "best" model was

$$\text{log salivary cotinine} = \beta_0 + \beta_1 \bullet \text{barning} + \beta_2 \bullet \text{topping} + \beta_3 \bullet \text{other} + \beta_4 \bullet \text{wet} + \beta_5 \bullet \text{lnnsmoke} + \beta_6 \bullet \text{barning} \bullet \text{smoke} + \beta_7 \bullet \text{topping} \bullet \text{smoke} + \beta_8 \bullet \text{other} \bullet \text{smoke}$$

Based on this model, the test for whether wet is significant is equivalent to testing H_0: $\beta_4=0$. We reject the null hypothesis (t-stat = 2.65, df = 1, p-value = 0.0083) and conclude that wet conditions have an effect on log salivary cotinine. Since there are no interactions, the differences lie only between workers in dry conditions and workers in wet conditions.

Based on this model, the test for whether task is significant is equivalent to testing H_0: $\beta_1=\beta_2=\beta_3=\beta_6=\beta_7=\beta_8=0$. We reject the null hypothesis (F-stat = 18.59, df=6, p-value < 0.0001) and conclude that task has an effect on cotinine level. Furthermore, to determine whether task is significant in regards to 2-way interactions, the test H_0: $\beta_6=\beta_7=\beta_8=0$ was performed. Again, we reject the null (F-stat = 20.27, df=3, p-value < 0.0001). To determine where the differences lie, contrasts were conducted looking at all pairwise comparisons of tasks. Bonferroni correction to p-values was utilized as it is more conservative than Scheffe and less ambiguous in this situation, however given the p-values obtained correction was unneccesary.

The contrasts are displayed in Table 6 below

Contrast	*DF*	*Contrast SS*	*Mean Square*	*F Value*	*Pr > F*
Barning Topping	1	47.7708168	47.7708168	37.40	<.0001
Barning Other	1	164.2166069	164.2166069	128.57	<.0001
Topping Other	1	57.3655073	57.3655073	44.91	<.0001

Table 6

For these contrasts, H_0: $\beta_i=\beta_j$ where the index corresponds to a task. For comparing barning, topping and other to the reference (priming), the t-values from the full model glm results were used (-5.84, -17.25 and -23.49 respectively, all p-values < 0.0001). As can be seen, the differences lies between each of the task categories.

In summary, cotinine levels are related to working task, wet conditions and cigarettes smoked. For tasks, priming contributes the most to cotinine, followed by barning, topping and other, in that order. Wet conditions also contribute to an increase in cotinine, but no interaction with task or smoking is observed. The greater the number of cigarettes smoked, the higher the cotinine levels observed, which is expected as cotinine is a nicotine metabolite. In addition though, an interaction between smoking and task is also observed, increasing the effect of each task with increasing number of cigarettes smoked.