# Effect coding versus dummy coding in analysis of data from factorial experiments

**Kari C. Kugler**
**Jessica B. Trail**
**John J. Dziak**
**Linda M. Collins**

*The Pennsylvania State University*

*Technical Report Series*
*#12-120*

*College of Health and Human Development*
*The Pennsylvania State University*

Address correspondence to Kari C. Kugler
The Methodology Center, The Pennsylvania State University
203 E. Calder Way, Suite 400, State College, PA 16801
Phone: (814) 863-9795
kck18@psu.edu

KEY WORDS: Factorial experiment, dummy coding, effect coding

**Abstract**

This technical report is intended to serve as a tutorial describing the differences between effect coding and dummy coding when the multiple regression approach is used to perform analysis of variance (ANOVA) with balanced (i.e., an equal number of subjects in each experimental condition) factorial designs. Using a hypothetical example of a $2^3$ factorial experiment, we present these two coding schemes for categorical independent variables and explain how the effects estimated can have different interpretations depending on the coding scheme used. Particular attention is paid to highlighting how and why differences exist.

## Introduction

The goal of this technical report is to explain something that is commonly written about in textbooks, but perhaps forgotten in practice: *how* a categorical independent variable is quantified in a regression analysis determines the interpretation of its associated regression coefficient. This report focuses on quantifying categorical variables when the regression approach is used to perform analysis of variance (ANOVA) on data from a balanced factorial experiment in which each factor has two levels. We compare and contrast the two most widely used approaches for quantifying categorical variables: dummy coding, in which the levels of the factor are coded 0 and 1, and effect coding, in which the levels of the factors are coded -1 and 1.

The two coding schemes always yield the same omnibus $F$. However, they can yield different estimates, test statistics, and $p$-values for the individual effects. This is because the individual effects estimated by the two coding schemes are different. In our view, neither of these coding schemes is right or wrong, but it is necessary to understand exactly what effects are being estimated in order to make an informed decision about whether to use dummy or effect coding. It is also helpful to understand what estimates are being provided by statistical packages.

The technical report is organized as follows. First we provide a brief introduction to the factorial design. We follow this with a description of the classical definitions of main effects and interactions. We then present the two different coding schemes for categorical variables. We provide a discussion of how to interpret regression coefficients when effect coding is used and when dummy coding is used. Next we provide a numerical example comparing effect and dummy coding. We then discuss how to analyze factorial experiments with SAS and SPSS. We conclude with a short discussion and summary of the take-away messages from this technical report.

**Brief Introduction to the Factorial Design**

Suppose an investigator wishes to examine the efficacy of three components of a smoking cessation intervention to increase self-efficacy after quitting (e.g., Baker et al., 2011; Collins et al., 2011): use of a nicotine patch; use of nicotine gum; and participation in counseling. The investigator decides to conduct a factorial experiment with the following three factors: (1) Nicotine Patch (Yes/No), (2) Gum (Yes/No), and (3) Counseling (Yes/No). This is a $2 \times 2 \times 2$, or $2^3$, factorial experiment. The design of this experiment is depicted in Table 1. (In the experimental condition column of Table 1, a "+" represents Yes and a "−" represents No.) As Table 1 shows, there are $2^3 = 8$ experimental conditions. Each experimental condition represents a unique combination of levels of the three factors, and all unique combinations are included in the design.

**Table 1. Experimental conditions for the hypothetical $2^3$ factorial experiment examining smoking cessation intervention components**

| Experimental condition number | Experimental condition | Factor 1 Nicotine Patch | Factor 2 Gum | Factor 3 Counseling |
|---|---|---|---|---|
| 1 | −−− | No | No | No |
| 2 | −−+ | No | No | Yes |
| 3 | −+− | No | Yes | No |
| 4 | −++ | No | Yes | Yes |
| 5 | +−− | Yes | No | No |
| 6 | +−+ | Yes | No | Yes |
| 7 | ++− | Yes | Yes | No |
| 8 | +++ | Yes | Yes | Yes |

Factorial experiments were originally developed to enable estimation of main effects and interaction effects in an efficient manner (Fisher, 1925). Referring to a $2^4$ factorial design, Fisher said,

> We have seen that the factorial arrangement possesses two advantages over experiments involving only single factors: (i) Greater *efficiency*, in that these factors are evaluated with the same precision by means of only a quarter of the number of

> observations that would otherwise be necessary; and (ii) Greater *comprehensiveness* in that, in addition to the 4 effects of single factors, their 11 possible interactions are evaluated. (p. 101)

Fisher goes on to list another, more subtle advantage of factorial experiments:

> There is a third advantage which, while less obvious than the former two, has an important bearing upon the utility of the experimental results in their practical application. This is that any conclusion…has a wider inductive basis when inferred from an experiment in which the quantities of other ingredients have been varied, than it would have from any amount of experimentation, in which these had been kept strictly constant… In fact, as the factorial arrangement well illustrates, we may, by deliberately varying in each case some of the conditions of the experiment, achieve a wider inductive basis for our conclusions, without in any degree impairing their precision. (Fisher, 1971, p. 102).

## Classical Definition of Effects

We remind the reader that the definitions of effects presented here pertain to factors with two levels (see Appendix A for a glossary of definitions).

**Main effects.** The classical definition of the main effect in a $2^k$ ANOVA (where $k =$ number of factors) is the difference between the mean response at one level of a particular factor and the mean response at the other level, collapsing over the levels of all remaining factors (Montgomery, 2009). For a $2^3$ factorial experiment with factors $A$, $B$, and $C$, the main effect of factor $A$ is represented by

$$\mathrm{ME}_A = \bar{\mu}_{(+\cdot\cdot)} - \bar{\mu}_{(-\cdot\cdot)},$$

where $\bar{\mu}_{(+\cdot\cdot)}$ and $\bar{\mu}_{(-\cdot\cdot)}$ represent the mean response for the "+" and "−" levels, respectively, of factor $A$, collapsing across the levels of factors $B$ and $C$ (the dot subscript means "summed over"). Here the notation $\bar{\mu}$ denotes a mean of means, that is, a mean taken across several experimental conditions. In the experiment depicted in Table 1, $\bar{\mu}_{(+\cdot\cdot)}$ is the mean response of experimental conditions 5-8 and $\bar{\mu}_{(-\cdot\cdot)}$ is the mean response of experimental conditions 1-4. Thus, in the factorial experiment described above, the main effect of Nicotine Patch is the difference between mean self-efficacy for Nicotine Patch at the Yes and No levels, collapsing over levels of Gum and Counseling.

Correspondingly,

$$\mathrm{ME}_B = \bar{\mu}_{(\cdot+\cdot)} - \bar{\mu}_{(\cdot-\cdot)},$$

and

$$\mathrm{ME}_c = \bar{\mu}_{(\cdot\cdot+)} - \bar{\mu}_{(\cdot\cdot-)}.$$

**Two-way interaction effects.** Conceptually, there is an $A \times B$ interaction if the effect of factor $A$ is different depending on the level of factor $B$. In a $2^k$ ANOVA, a two-way interaction is the average of the difference in the effect of a particular factor across the levels of a second factor, collapsing over all other factors (Montgomery, 2009; p.217). For example, in a $2^3$ ANOVA, this is represented by the following:

$$\mathrm{INT}_{A\times B} = \frac{1}{2}\left[\left(\bar{\mu}_{(++\cdot)} - \bar{\mu}_{(-+\cdot)}\right) - \left(\bar{\mu}_{(+-\cdot)} - \bar{\mu}_{(--\cdot)}\right)\right].$$

In the example experiment, the Nicotine Patch x Gum interaction is the average of the difference in the effect of Nicotine Patch on self-efficacy when Gum = Yes and the effect of Nicotine Patch when Gum = No, collapsed over levels of Counseling. If the effect of Nicotine Patch on self-efficacy is the same when Gum=Yes and when Gum=No, there is no Nicotine Patch x Gum interaction, and this difference is zero.

**Three-way interaction effects**. Conceptually, there is an $A \times B \times C$ interaction when the $A \times B$ interaction effect is different depending on the level of factor $C$. In a $2^k$ ANOVA, a three-way interaction is the average of the difference in the two-way interaction effects at differing levels of a third factor, collapsing over any remaining factors. For example, in a $2^3$ ANOVA, this is represented by the following:

$$\text{INT}_{A\times B\times C}=\frac{1}{2}\left[\tfrac{1}{2}\left[\left(\mu_{(+++)}-\mu_{(-++)}\right)-\left(\mu_{(+-+)}-\mu_{(--+)}\right)\right]-\tfrac{1}{2}\left[\left(\mu_{(++-)}-\mu_{(-+-)}\right)-\left(\mu_{(+--)}-\mu_{(---)}\right)\right]\right].$$

The Nicotine Patch × Gum × Counseling interaction is the average of the difference between the Nicotine Patch × Gum interaction when Counseling = Yes and the Nicotine Patch × Gum interaction when Counseling = No. If there were additional factors in the experiment, this difference would be collapsed over those factors as well.

## Coding Schemes for Categorical Variables

In the section that follows, we present the details of dummy coding and effect coding. In later sections we demonstrate how these two coding schemes differ, how the effects they estimate are different, and how one could come to a different conclusion depending on the type of coding scheme used. In particular, we will demonstrate that effect codes produce the classically defined effects, but dummy codes produce different effects. We begin with dummy coding, which is the type of coding perhaps most familiar to behavioral researchers.

**Dummy coding.** In a dummy coding system, 1s and 0s are used, where 1 represents membership in a level/category and 0 represents non-membership. To represent a variable with $K$ categories, $K - 1$ dummy variables are needed. In our hypothetical example, the variable Nicotine Patch has only two categories, Yes and No. Therefore, a single dummy variable is needed to represent that variable; for participants who received the Nicotine Patch the dummy variable has a value of "1," and for those who did not receive the Nicotine Patch the dummy variable has a value of "0." The same approach would be used for Gum and Counseling. (Readers interested in learning more about dummy coding for variables with more than two categories are referred to Hardy, 1993, pages 8-9).

Table 2 shows the dummy coding scheme for the factorial design in Table 1. The rows correspond to the experimental conditions and the columns correspond to the dummy variables for the individual factors and their product terms. We are avoiding the use of the terms "main effect" and "interaction" here, because, as will be demonstrated below, the effects estimated when dummy codes are used are not main effects and interactions according to the classical definitions above.

The column labeled $Z_{NP}$ represents the dummy variable for the factor Nicotine Patch (the letter $Z$ was arbitrarily chosen and should not be confused with a $Z$ score). All those at the No level are given a 0 (Rows 1-4), while all those at the Yes level are given a 1 (Rows 5-8). The column labeled $Z_{GUM}$ represents the dummy variable for the factor Gum. Again, a 0 is listed for all those at the No level (Rows 1,2,5,6); a 1 is listed for all those at the Yes level (Rows 3,4,7,8). The column $Z_{CS}$ represents the dummy variable for the factor Counseling. Here, Rows 1,3,5,7 are given a 0 for the No level, and Rows 2,4,6,8 are given a 1 for the Yes level.

The remaining columns represent the variables for the product terms and are constructed by multiplying the elements in the vectors relating to the individual factors involved in the product. For example, $Z_{NP \times GUM}$ represents the Nicotine Patch × Gum product and is created by multiplying the elements in the vector for $Z_{NP}$ by the elements in the vector for $Z_{GUM}$.

**Table 2. Dummy coding scheme for the eight experimental conditions**

| Experimental condition number | Experimental condition | Dummy codes | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | $Z_{NP}$ | $Z_{GUM}$ | $Z_{CS}$ | $Z_{NP \times GUM}$ | $Z_{NP \times CS}$ | $Z_{GUM \times CS}$ | $Z_{NP \times GUM \times CS}$ |
| 1 | − − − | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | − − + | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| 3 | − + − | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 4 | − + + | 0 | 1 | 1 | 0 | 0 | 1 | 0 |
| 5 | + − − | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6 | + − + | 1 | 0 | 1 | 0 | 1 | 0 | 0 |
| 7 | + + − | 1 | 1 | 0 | 1 | 0 | 0 | 0 |
| 8 | + + + | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

**Effect coding.** Effect coding is similar to dummy coding in that a 1 represents membership in a level/category. However, in effect coding for two-level variables, non-membership is coded -1 (rather than 0; see Hardy, 1993, pages 64-71, for a description of effect coding for variables with more than two levels). As in dummy coding, the variables representing the product terms in effect coding are constructed by multiplying the elements in the vectors corresponding to the factors involved in the product. Table 3 presents the effect coding scheme for the hypothetical factorial experiment in Table 1.

**Table 3. Effect coding scheme for the eight experimental conditions**

| Experimental condition number | Experimental condition | Effect codes | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | $X_{NP}$ | $X_{GUM}$ | $X_{CS}$ | $X_{NP\times GUM}$ | $X_{NP\times CS}$ | $X_{GUM\times CS}$ | $X_{NP\times GUM\times CS}$ |
| 1 | − − − | −1 | −1 | −1 | 1 | 1 | 1 | −1 |
| 2 | − − + | −1 | −1 | 1 | 1 | −1 | −1 | 1 |
| 3 | − + − | −1 | 1 | −1 | −1 | 1 | −1 | 1 |
| 4 | − + + | −1 | 1 | 1 | −1 | −1 | 1 | −1 |
| 5 | + − − | 1 | −1 | −1 | −1 | −1 | 1 | 1 |
| 6 | + − + | 1 | −1 | 1 | −1 | 1 | −1 | −1 |
| 7 | + + − | 1 | 1 | −1 | 1 | −1 | −1 | −1 |
| 8 | + + + | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

**Interpretation of Regression Coefficients in Regression Models**

**Population means for each experimental condition when effect coding is used.** Below is the hypothetical experiment with three factors, each with two levels, expressed as an effect-coded regression equation:

$$\mu = \beta_0 + \beta_1 X_{NP} + \beta_2 X_{GUM} + \beta_3 X_{CS} + \beta_{12} X_{NP\times GUM} + \beta_{13} X_{NP\times CS} + \beta_{23} X_{GUM\times CS} + \beta_{123} X_{NP\times GUM\times CS}. \quad (1)$$

In Equation (1) above, $\mu$ represents the population mean of the outcome of interest, that is, self-efficacy after quitting, and each $X$ corresponds to a column from Table 3. Each of the regression coefficients ($\beta_1$, $\beta_2$, $\beta_3$, etc.) expresses the expected change in $\mu$ given a one unit change in the associated *X,* holding all else constant.

Table 4 shows the population means for each of the experimental conditions expressed in terms of Equation (1) using effect coding. For experimental condition 1, in which each participant is at the No (-1) level for the three factors, Equation (1) is

$$\beta_0 + \beta_1(-1) + \beta_2(-1) + \beta_3(-1) + \beta_{12}(-1)(-1) + \beta_{13}(-1)(-1) + \beta_{23}(-1)(-1) + \beta_{123}(-1)(-1)(-1).$$

The population mean for this experimental condition, $\mu_{(---)}$, is shown in Row 1 of Table 4.

For experimental condition 8, in which each participant is at the Yes (+) level for all the three factors, Equation (1) is

$$\beta_0 + \beta_1(1) + \beta_2(1) + \beta_3(1) + \beta_{12}(1)(1) + \beta_{13}(1)(1) + \beta_{23}(1)(1) + \beta_{123}(1)(1)(1).$$

The population mean for this treatment condition, $\mu_{(+++)}$, is shown in Row 8 of Table 4.

When effect coding is used, the intercept, $\beta_0$, is the grand mean. This can be seen by examining Table 4 carefully. Assuming a balanced design, the grand mean is the average of all the experimental condition means. Expressing these means in terms of regression coefficients and taking the average shows that the result is $\beta_0$.

**Table 4. Population means for each experimental condition when effect coding is used**

| Experimental condition number | Experimental condition | Population means |
|---|---|---|
| 1 | − − − | $\mu_{(---)} = \beta_0 - \beta_1 - \beta_2 - \beta_3 + \beta_{12} + \beta_{13} + \beta_{23} - \beta_{123}$ |
| 2 | − − + | $\mu_{(--+)} = \beta_0 - \beta_1 - \beta_2 + \beta_3 + \beta_{12} - \beta_{13} - \beta_{23} + \beta_{123}$ |
| 3 | − + − | $\mu_{(-+-)} = \beta_0 - \beta_1 + \beta_2 - \beta_3 - \beta_{12} + \beta_{13} - \beta_{23} + \beta_{123}$ |
| 4 | − + + | $\mu_{(-++)} = \beta_0 - \beta_1 + \beta_2 + \beta_3 - \beta_{12} - \beta_{13} + \beta_{23} - \beta_{123}$ |
| 5 | + − − | $\mu_{(+--)} = \beta_0 + \beta_1 - \beta_2 - \beta_3 - \beta_{12} - \beta_{13} + \beta_{23} + \beta_{123}$ |
| 6 | + − + | $\mu_{(+-+)} = \beta_0 + \beta_1 - \beta_2 + \beta_3 - \beta_{12} + \beta_{13} - \beta_{23} - \beta_{123}$ |
| 7 | + + − | $\mu_{(++-)} = \beta_0 + \beta_1 + \beta_2 - \beta_3 + \beta_{12} - \beta_{13} - \beta_{23} - \beta_{123}$ |
| 8 | + + + | $\mu_{(+++)} = \beta_0 + \beta_1 + \beta_2 + \beta_3 + \beta_{12} + \beta_{13} + \beta_{23} + \beta_{123}$ |

**Interpretation of effects in effect coding.**

***Main effects.*** The main effect of Nicotine Patch is computed by taking the difference in the average of Rows 5-8 (Nicotine Patch = Yes) and average of Rows 1-4 (Nicotine Patch = No).

$$\bar{\mu}_{(+\cdot\cdot)} = \frac{1}{4}\Big[(\beta_0+\beta_1-\beta_2-\beta_3-\beta_{12}-\beta_{13}+\beta_{23}+\beta_{123})+(\beta_0+\beta_1-\beta_2+\beta_3-\beta_{12}+\beta_{13}-\beta_{23}-\beta_{123})$$
$$+(\beta_0+\beta_1+\beta_2-\beta_3+\beta_{12}-\beta_{13}-\beta_{23}-\beta_{123})+(\beta_0+\beta_1+\beta_2+\beta_3+\beta_{12}+\beta_{13}+\beta_{23}+\beta_{123})\Big]$$
$$= \frac{1}{4}\big[4\beta_0+4\beta_1\big] = \beta_0+\beta_1$$
$$\bar{\mu}_{(-\cdot\cdot)} = \frac{1}{4}\big[4\beta_0-4\beta_1\big] = \beta_0-\beta_1$$

Using the population means for each experimental condition in Table 4, the main effect is expressed as follows:

$$\mathrm{ME}_{NP} = \bar{\mu}_{(+\cdot\cdot)} - \bar{\mu}_{(-\cdot\cdot)} = (\beta_0+\beta_1)-(\beta_0-\beta_1) = 2\beta_1$$

$$\beta_1 = \frac{1}{2} * \mathrm{ME}_{NP}$$

As shown in this equation, with effect coding the $\beta$s correspond to the main effects according to the classical definition. In order to compute the actual main effect, a simple linear transformation is required that is, multiplying the regression coefficient by a scaling constant of 2.

*Two-way interaction effects.* Using the same approach, the effect of the Nicotine Patch × Gum interaction is calculated as follows:

$$\bar{\mu}_{(++\cdot)} = \frac{1}{2}\big[2\beta_0+2\beta_1+2\beta_2+2\beta_{12}\big]$$
$$\bar{\mu}_{(-+\cdot)} = \frac{1}{2}\big[2\beta_0-2\beta_1+2\beta_2-2\beta_{12}\big]$$
$$\bar{\mu}_{(+-\cdot)} = \frac{1}{2}\big[2\beta_0+2\beta_1-2\beta_2-2\beta_{12}\big]$$
$$\bar{\mu}_{(--\cdot)} = \frac{1}{2}\big[2\beta_0-2\beta_1-2\beta_2+2\beta_{12}\big]$$
$$\mathrm{INT}_{NP\times GUM} = \frac{1}{2}\big[(2\beta_1+2\beta_{12})-(2\beta_1-2\beta_{12})\big] = 2\beta_{12}$$
$$\beta_{12} = \frac{1}{2}\mathrm{INT}_{NP\times GUM}$$

Thus, with effect coding, $\beta_{12}$ represents one-half the classical definition of a two-way interaction. In other words, the effect of the Nicotine Patch × Gum interaction is simply $2\beta_{12}$.

*Three-way interaction effects.* Again, using the same approach, the following represents the Nicotine Patch × Gum × Counseling interaction effect:

$$\begin{aligned}\mathrm{INT}_{NP\times GUM\times CS} = \frac{1}{2}\Big[ & \tfrac{1}{2}\big[(2\beta_1 + 2\beta_{12} + 2\beta_{13} + 2\beta_{123}) - (2\beta_1 - 2\beta_{12} + 2\beta_{13} - 2\beta_{123})\big] \\ & - \tfrac{1}{2}\big[(2\beta_1 + 2\beta_{12} - 2\beta_{13} - 2\beta_{123}) - (2\beta_1 - 2\beta_{12} - 2\beta_{13} + 2\beta_{123})\big]\Big] = 2\beta_{123}\end{aligned}$$

$$\beta_{123} = \frac{1}{2}\mathrm{INT}_{NP\times GUM\times CS}$$

Similar to all the other effects in a model with effect coding, the regression coefficient, $\beta_{123}$, represents one-half the classical definition of a three-way interaction for our experiment.

**Summary: Effect-coded effects.** When effect coding is used, the regression coefficients in a regression model are equivalent to the classically defined main effects and interactions, except for a scaling constant (that does not affect hypothesis tests). A simple multiplication of the regression coefficient by the scaling constant of 2 produces the estimated effects for the special case of two-level factors.

**Population means for each experimental condition when dummy coding is used.** Below is the hypothetical experiment with three factors, each with two levels, now expressed as a regression equation using dummy coding:

$$\mu = \alpha_0 + \alpha_1 Z_{NP} + \alpha_2 Z_{GUM} + \alpha_3 Z_{CS} + \alpha_{12} Z_{NP\times GUM} + \alpha_{13} Z_{NP\times CS} + \alpha_{23} Z_{GUM\times CS} + \alpha_{123} Z_{NP\times GUM\times CS}. \quad (2)$$

In Equation (2), $\mu$ again represents the outcome of interest, that is, self-efficacy after quitting. Each $Z$ corresponds to a column from Table 2. We use $\alpha$ to represent regression coefficients to distinguish these from the regression coefficients in Equation 1. The intercept, $\alpha_0$, is the mean of $\mu$ when all of the independent variables are set to zero, in this case, the No level of each factor. Each of the remaining coefficients $(\alpha_1, \alpha_2, \alpha_3, \text{etc.})$ represents the expected change in $\mu$ given a one unit change in the associated $Z$, holding all else constant.

Table 5 shows the population means for each of the experimental conditions expressed in terms of Equation (2) when dummy coding is used. For example, the expected mean for experimental condition 1, in which each of the three factors is set to the No (−) level is

$$\mu_{(---)} = \alpha_0 + \alpha_1(0) + \alpha_2(0) + \alpha_3(0) + \alpha_{12}(0)(0) + \alpha_{13}(0)(0) + \alpha_{23}(0)(0) + \alpha_{123}(0)(0)(0) = \alpha_0,$$

as shown in Row 1 of Table 5. The population mean for experimental condition 8 is

$$\begin{aligned}\mu_{(+++)} &= \alpha_0 + \alpha_1(1) + \alpha_2(1) + \alpha_3(1) + \alpha_{12}(1)(1) + \alpha_{13}(1)(1) + \alpha_{23}(1)(1) + \alpha_{123}(1)(1)(1)\\ &= \alpha_0 + \alpha_1 + \alpha_2 + \alpha_3 + \alpha_{12} + \alpha_{13} + \alpha_{23} + \alpha_{123},\end{aligned}$$

as shown in Row 8 of Table 5.

**Table 5. Population means for each experimental condition when dummy coding is used**

| Experimental condition number | Experimental condition | Population means |
|---|---|---|
| 1 | − − − | $\mu_{(---)} = \alpha_0$ |
| 2 | − − + | $\mu_{(--+)} = \alpha_0 + \alpha_3$ |
| 3 | − + − | $\mu_{(-+-)} = \alpha_0 + \alpha_2$ |
| 4 | − + + | $\mu_{(-++)} = \alpha_0 + \alpha_2 + \alpha_3 + \alpha_{23}$ |
| 5 | + − − | $\mu_{(+--)} = \alpha_0 + \alpha_1$ |
| 6 | + − + | $\mu_{(+-+)} = \alpha_0 + \alpha_1 + \alpha_3 + \alpha_{13}$ |
| 7 | + + − | $\mu_{(++-)} = \alpha_0 + \alpha_1 + \alpha_2 + \alpha_{12}$ |
| 8 | + + + | $\mu_{(+++)} = \alpha_0 + \alpha_1 + \alpha_2 + \alpha_3 + \alpha_{12} + \alpha_{13} + \alpha_{23} + \alpha_{123}$ |

Comparison of Tables 4 and 5shows that the type of coding scheme used produces dramatically different parameterizations of the population means for each of the experimental conditions. It is also worth noting that the intercept has a different meaning depending on the type of coding scheme used. Recall that when effect coding is used, the intercept, $\beta_0$, is the grand mean of $\mu$. In contrast, as Table 5 shows, the intercept in a dummy-coded regression model, $\alpha_0$, is the mean of $\mu$ when each of the three factors is set to the No (−) level.

**Expressing dummy-coded effects in terms of classical effects.** When dummy coding is used and product terms are included in the model, the regression coefficients do not correspond to the main effects and interactions according to the classical definition (Hardy, 1993). This is not inherently bad or a mistake: the effects that are estimated may be of interest in a particular study. As will be shown later, these effects are referred to as "simple effects," that is, an effect of a factor at one level of another factor(s). In this technical report we will call the regression coefficients corresponding to the dummy-coded variables of a single factor as "first-order effects" (1OE) to distinguish them from classically defined main effects. Similarly, we will refer to the regression coefficients corresponding to the dummy-coded variables of two-factor product terms as "second-order effects" (2OE), and to the regression coefficients corresponding to the dummy-coded variables of three-factor product terms as "third-order effects" (3OE), to distinguish them from classically defined interactions.

***First-order effects.*** As discussed above, the main effect of the Nicotine Patch is computed by taking the difference between the average of Rows 5-8 (Nicotine Patch = Yes)

$$\begin{aligned}\bar{\mu}_{(+\cdot\cdot)} &= \frac{1}{4}\Big[(\alpha_0 + \alpha_1) + (\alpha_0 + \alpha_1 + \alpha_3 + \alpha_{13}) + (\alpha_0 + \alpha_1 + \alpha_2 + \alpha_{12}) \\ &\quad + (\alpha_0 + \alpha_1 + \alpha_2 + \alpha_3 + \alpha_{12} + \alpha_{13} + \alpha_{23} + \alpha_{123})\Big] \\ &= \frac{1}{4}\Big[4\alpha_0 + 4\alpha_1 + 2\alpha_2 + 2\alpha_3 + 2\alpha_{12} + 2\alpha_{13} + \alpha_{23} + \alpha_{123}\Big]\end{aligned}$$

and Rows 1-4 (Nicotine Patch = No)

$$\begin{aligned}\bar{\mu}_{(-\cdot\cdot)} &= \frac{1}{4}\Big[(\alpha_0) + (\alpha_0 + \alpha_3) + (\alpha_0 + \alpha_2) + (\alpha_0 + \alpha_2 + \alpha_3 + \alpha_{23})\Big] \\ &= \frac{1}{4}[4\alpha_0 + 2\alpha_2 + 2\alpha_3 + \alpha_{23}]\end{aligned}$$

$$\mathrm{ME}_{NP} = \bar{\mu}_{(+\cdot\cdot)} - \bar{\mu}_{(-\cdot\cdot)} = \frac{1}{4}\Big[4\alpha_1 + 2\alpha_{12} + 2\alpha_{13} + \alpha_{123}\Big] = \alpha_1 + \frac{1}{2}\alpha_{12} + \frac{1}{2}\alpha_{13} + \frac{1}{4}\alpha_{123}.$$

Based on this, it is now possible to solve for $\alpha_1$:

$$\alpha_1 = 1\mathrm{OE}_{NP} = \mathrm{ME}_{NP} - \mathrm{INT}_{NP\times GUM} - \mathrm{INT}_{NP\times CS} + \mathrm{INT}_{NP\times GUM\times CS}$$

As shown in this equation, when dummy coding is used $\alpha_1$ does not represent the main effect of Nicotine Patch according to the classical definition. Instead, $\alpha_1$ is a linear combination of the classically defined main effect and the higher-order effects that include this factor, in this case two two-way interactions and the three-way interaction. Thus, a test of the null hypothesis that $\alpha_1$ equals zero is not a test of the null hypothesis that the main effect of Nicotine Patch equals zero.

This raises the question of how $\alpha_1$ should be interpreted. $\alpha_1$ is the effect of Nicotine Patch when all the other factors in the model are set to zero. An explanation of this is provided in Appendix B.

***Second-order effects.*** Using the same approach as above, the Nicotine Patch × Gum interaction is calculated as follows:

$$\bar{\mu}_{(++\cdot)} = \frac{1}{2}\left[2\alpha_0 + 2\alpha_1 + 2\alpha_2 + \alpha_3 + 2\alpha_{12} + \alpha_{13} + \alpha_{23} + \alpha_{123}\right]$$

$$\bar{\mu}_{(-+\cdot)} = \frac{1}{2}\left[2\alpha_0 + 2\alpha_2 + \alpha_3 + \alpha_{23}\right]$$

$$\bar{\mu}_{(+-\cdot)} = \frac{1}{2}\left[2\alpha_0 + 2\alpha_1 + \alpha_3 + \alpha_{13}\right]$$

$$\bar{\mu}_{(--\cdot)} = \frac{1}{2}[2\alpha_0 + \alpha_3]$$

$$\mathrm{INT}_{NP\times GUM} = \frac{1}{2}\left[\tfrac{1}{2}\left[(2\alpha_1 + 2\alpha_{12} + \alpha_{13} + \alpha_{123}) - (2\alpha_1 + \alpha_{13})\right]\right] = \frac{1}{2}\alpha_{12} + \frac{1}{4}\alpha_{123}$$

$$\alpha_{12} = 2\mathrm{OE}_{NP\times GUM} = 2 * \mathrm{INT}_{NP\times GUM} - 2 * \mathrm{INT}_{NP\times GUM\times CS}$$

Thus the test of the null hypothesis that $\alpha_{12}$ equals zero is not a test of the null hypothesis that the Nicotine Patch × Gum interaction equals 0 (Chakraborty, Collins, Strecher, & Murphy, 2009). Instead, $\alpha_{12}$ is a linear combination of the classically defined two-way interaction and the three-way interaction.

Appendix B shows that the hypothesis test should be interpreted as a test of the null hypothesis that the Nicotine Patch × Gum interaction equals 0 when Counseling is set to zero. The same holds for the other second-order effects in the model, i.e., Nicotine Patch × Counseling (e.g., $\alpha_{13}$) and Gum × Counseling (e.g., $\alpha_{23}$).

***Third-order effects.*** Using the same approach, the following is the Nicotine Patch × Gum × Counseling interaction effect:

$$\mathrm{INT}_{NP\times GUM\times CS} = \frac{1}{4}\Big[\big[(\alpha_1+\alpha_{12}+\alpha_{13}+\alpha_{123})-(\alpha_1+\alpha_{13})\big]-\big[(\alpha_1+\alpha_{12})-(\alpha_1)\big]\Big] = \frac{1}{4}\alpha_{123}$$

$$\alpha_{123} = 3\mathrm{OE}_{NP\times GUM\times CS} = 4*\mathrm{INT}_{NP\times GUM\times CS}$$

This third-order effect is the only effect in this model that corresponds to the classically defined effect, in this case a three-way interaction. <u>In general, the test of significance for the highest-order effect is equivalent in dummy coding and effect coding</u>.

**Summary: Dummy-coding effects.** When dummy coding is used (and provided that product terms are included in the model), the regression coefficients no longer estimate effects corresponding to the classically defined main effects or interactions; rather, they estimate simple effects. The one exception is the highest order effect, which is equivalent to the classically defined effect. Table 6 provides a summary of the effects with effect coding and dummy coding.

**Table 6. Summary table of effects with effect coding and dummy coding**

| Effect coding | | | Dummy coding | | |
|---|---|---|---|---|---|
| Regression coefficient | Regression coefficient expressed in terms of experimental condition means | Definition of regression coefficient | Regression coefficient | Regression coefficient expressed in terms of experimental condition means | Definition of regression coefficient (see Appendix B) |
| $\beta_1$ | $\bar{\mu}_{(+\cdot\cdot)} - \bar{\mu}_{(-\cdot\cdot)}$ | ½ the main effect of Nicotine Patch | $\alpha_1$ | $\mu_{(+--)} - \mu_{(---)}$ | The effect of Nicotine Patch, when Gum & Counseling are 0 |
| $\beta_2$ | $\bar{\mu}_{(\cdot+\cdot)} - \bar{\mu}_{(\cdot-\cdot)}$ | ½ the main effect of Gum | $\alpha_2$ | $\mu_{(-+-)} - \mu_{(---)}$ | The effect of Gum, when Nicotine Patch & Counseling are 0 |
| $\beta_3$ | $\bar{\mu}_{(\cdot\cdot+)} - \bar{\mu}_{(\cdot\cdot-)}$ | ½ the main effect of Counseling | $\alpha_3$ | $\mu_{(--+)} - \mu_{(---)}$ | The effect of Counseling, when Nicotine Patch & Gum are 0 |
| $\beta_{12}$ | $\frac{1}{4}\left[\left(\bar{\mu}_{(++\cdot)} - \bar{\mu}_{(-+\cdot)}\right) - \left(\bar{\mu}_{(+-\cdot)} - \bar{\mu}_{(--\cdot)}\right)\right]$ | ½ the interaction effect of Nicotine Patch × Gum | $\alpha_{12}$ | $\left(\mu_{(++-)} - \mu_{(-+-)}\right) - \left(\mu_{(+--)} - \mu_{(---)}\right)$ | The effect of Nicotine Patch × Gum when Counseling is 0 |
| $\beta_{13}$ | $\frac{1}{4}\left[\left(\bar{\mu}_{(+\cdot+)} - \bar{\mu}_{(-\cdot+)}\right) - \left(\bar{\mu}_{(+\cdot-)} - \bar{\mu}_{(-\cdot-)}\right)\right]$ | ½ the interaction effect of Nicotine Patch × Counseling | $\alpha_{13}$ | $\left(\mu_{(+-+)} - \mu_{(--+)}\right) - \left(\mu_{(+--)} - \mu_{(---)}\right)$ | The effect of Nicotine Patch × Counseling when Gum is 0 |
| $\beta_{23}$ | $\frac{1}{4}\left[\left(\bar{\mu}_{(\cdot++)} - \bar{\mu}_{(\cdot-+)}\right) - \left(\bar{\mu}_{(\cdot+-)} - \bar{\mu}_{(\cdot--)}\right)\right]$ | ½ the interaction effect of Gum × Counseling | $\alpha_{23}$ | $\left(\mu_{(-++)} - \mu_{(--+)}\right) - \left(\mu_{(-+-)} - \mu_{(---)}\right)$ | The effect of Gum × Counseling when Nicotine Patch is 0 |
| $\beta_{123}$ | $\frac{1}{8}\left[\left[\left(\mu_{(+++)} - \mu_{(-++)}\right) - \left(\mu_{(+-+)} - \mu_{(--+)}\right)\right] - \left[\left(\mu_{(++-)} - \mu_{(-+-)}\right) - \left(\mu_{(+--)} - \mu_{(---)}\right)\right]\right]$ | ½ the interaction effect of Nicotine Patch × Gum × Counseling | $\alpha_{123}$ | $\left[(\mu_{(+++)} - \mu_{(-++)}) - (\mu_{(+-+)} - \mu_{(--+)})\right] - \left[(\mu_{(++-)} - \mu_{(-+-)}) - (\mu_{(+--)} - \mu_{(---)})\right]$ | 4× the interaction effect of Nicotine Patch × Gum × Counseling |

## A Numerical Example

**Study design.** To illustrate dummy coding and effect coding, we present a hypothetical numerical example of the aforementioned factorial experiment.

Recall that we are interested in examining the following research questions:

a) Does the use of a Nicotine Patch alter self-efficacy compared to not using a Nicotine Patch, collapsing over levels of Gum and Counseling?

   i. $H_0 : \bar{\mu}_{(+\cdot\cdot)} - \bar{\mu}_{(-\cdot\cdot)} = 0$ vs.

   ii. $H_A : \bar{\mu}_{(+\cdot\cdot)} - \bar{\mu}_{(-\cdot\cdot)} \neq 0$

b) Does the use of Gum alter self-efficacy compared to not using Gum, collapsing over levels of Nicotine Patch and Counseling?

   i. $H_0 : \bar{\mu}_{(\cdot+\cdot)} - \bar{\mu}_{(\cdot-\cdot)} = 0$ vs.

   ii. $H_A : \bar{\mu}_{(\cdot+\cdot)} - \bar{\mu}_{(\cdot-\cdot)} \neq 0$

c) Does receiving Counseling alter self-efficacy compared to not receiving Counseling, collapsing over levels of Nicotine Patch and Gum?

   i. $H_0 : \bar{\mu}_{(\cdot\cdot+)} - \bar{\mu}_{(\cdot\cdot-)} = 0$ vs.

   ii. $H_A : \bar{\mu}_{(\cdot\cdot+)} - \bar{\mu}_{(\cdot\cdot-)} \neq 0$

d) Does the effect of the Nicotine Patch on self-efficacy vary based on whether or not Gum was administered, collapsing over levels of Counseling?

   i. $H_0 : \frac{1}{2}\left[\left(\bar{\mu}_{(++\cdot)} - \bar{\mu}_{(-+\cdot)}\right) - \left(\bar{\mu}_{(+-\cdot)} - \bar{\mu}_{(--\cdot)}\right)\right] = 0$ vs.

   ii. $H_A : \frac{1}{2}\left[\left(\bar{\mu}_{(++\cdot)} - \bar{\mu}_{(-+\cdot)}\right) - \left(\bar{\mu}_{(+-\cdot)} - \bar{\mu}_{(--\cdot)}\right)\right] \neq 0$

e) Does the effect of the Nicotine Patch on self-efficacy vary based on whether or not participants received Counseling, collapsing over levels of Gum?

i. $H_0: \frac{1}{2}\left[\left(\bar{\mu}_{(+\cdot+)} - \bar{\mu}_{(-\cdot+)}\right) - \left(\bar{\mu}_{(+\cdot-)} - \bar{\mu}_{(-\cdot-)}\right)\right] = 0$ vs.

ii. $H_A: \frac{1}{2}\left[\left(\bar{\mu}_{(+\cdot+)} - \bar{\mu}_{(-\cdot+)}\right) - \left(\bar{\mu}_{(+\cdot-)} - \bar{\mu}_{(-\cdot-)}\right)\right] \neq 0$

f) Does the effect of Gum on self-efficacy vary based on whether or not participants received Counseling, collapsing over levels of Nicotine Patch?

i. $H_0: \frac{1}{2}\left[\left(\bar{\mu}_{(\cdot++)} - \bar{\mu}_{(\cdot-+)}\right) - \left(\bar{\mu}_{(\cdot+-)} - \bar{\mu}_{(\cdot--)}\right)\right] = 0$ vs.

ii. $H_A: \frac{1}{2}\left[\left(\bar{\mu}_{(\cdot++)} - \bar{\mu}_{(\cdot-+)}\right) - \left(\bar{\mu}_{(\cdot+-)} - \bar{\mu}_{(\cdot--)}\right)\right] \neq 0$

g) Does the interaction between Nicotine Patch and Gum vary based on whether or not participants received Counseling?

i. $H_0: \frac{1}{4}\left[\mu_{(+++)} - \mu_{(-++)} - \mu_{(+-+)} + \mu_{(--+)} - \mu_{(++-)} + \mu_{(-+-)} + \mu_{(+--)} - \mu_{(---)}\right] = 0$ vs.

ii. $H_A: \frac{1}{4}\left[\mu_{(+++)} - \mu_{(-++)} - \mu_{(+-+)} + \mu_{(--+)} - \mu_{(++-)} + \mu_{(-+-)} + \mu_{(+--)} - \mu_{(---)}\right] \neq 0$

For this experiment, 32 young adults are randomly assigned to one of the eight experimental conditions and followed over time. Below are the results of the hypothetical experiment, presented first for effect-coded independent variables and then for dummy-coded independent variables. We assume balanced data (i.e., there are an equal number of subjects in each condition at the end of our study).

Please note that the numerical example is based on artificial data.

**Effect coding results and interpretation.** Table 7 shows the output for a single multiple regression model that includes the three main effects, three two-way interaction effects, and one three-way interaction effect using PROC REG procedure in SAS/STAT software (SAS version 9.2, SAS Institute Inc.).

**Table 7. Annotated regression output for effect coded variables**

The REG Procedure

Model: MODEL1
Dependent Variable: Y

Number of Observations Read 32
Number of Observations Used 32

Analysis of Variance

| Source | DF | Sum of Squares | Mean Square | F Value | Pr>F |
|---|---|---|---|---|---|
| Model | 7 | 1686.53173 | 240.93310 | 217.32 | <.0001 |
| Error | 24 | 26.60724 | 1.10863 | | |
| Corrected Total | 31 | 1713.13897 | | | |

| | |
|---|---|
| Root MSE | 1.05292 |
| R-Square | 0.9845 |
| Dependent Mean | -0.17256 |
| Adj R-Sq | 0.9799 |
| Coeff Var | -610.15872 |

| Variable | DF | Parameter Estimate | Standard Error | t Value | Pr>1 |
|---|---|---|---|---|---|
| Intercept | 1 | -0.17256 | 0.18613 | -0.93 | 0.3631 |
| $X_{NP}$ | 1 | 6.51605 | 0.18613 | 35.01 | <.0001 |
| $X_{GUM}$ | 1 | 0.11797 | 0.18613 | 0.63 | 0.5322 |
| $X_{CS}$ | 1 | -0.21210 | 0.18613 | -1.14 | 0.2657 |
| $X_{NPXGUM}$ | 1 | 3.13873 | 0.18613 | 16.86 | <.0001 |
| $X_{NPXCS}$ | 1 | 0.40815 | 0.18613 | 2.19 | 0.0383 |
| $X_{GUMXCS}$ | 1 | -0.31576 | 0.18613 | -1.70 | 0.1027 |
| $X_{NPXGUMXCS}$ | 1 | 0.26153 | 0.18613 | 1.41 | 0.1728 |

Most researchers begin to interpret the results by looking at the overall $F$ statistic for the model. In this example, the omnibus $F$ statistic is significant ($F$=217.32, $p$<0.01), meaning that at least one of the effects in the model is not equal to zero. Next, attention is drawn to the parameter estimates. Beginning with the regression coefficient for $X_{NP}$, the parameter estimate of 6.52 is significant ($p$<0.01), meaning that, on average, the use of Nicotine Patch increases self-efficacy by 13.04 units (6.52*2) compared to no use of the Nicotine Patch, *collapsing over levels of Gum and Counseling*. The parameter estimates for the Gum and Counseling regression coefficients are not significant, $p$=0.53 and $p$=0.27, respectively, meaning that there is not enough evidence to suggest that these factors have an effect on self-efficacy on average in these artificial data.

There are two significant two-way interactions: the Nicotine Patch × Gum interaction ($p$<0.01) and Nicotine Patch × Counseling interaction ($p$=0.04). We can interpret the regression coefficient for the Nicotine Patch × Gum interaction as follows: The expected difference in the effect of Nicotine Patch on self-efficacy when Gum is used and the effect of Nicotine Patch on craving when Gum is not used, collapsing over levels of Counseling, is 6.28 units (3.14*2). In other words, the effect of Nicotine Patch on self-efficacy increases by 6.28 units when Gum is used compared to when Gum is not used, collapsing over levels of Counseling.

**Dummy coding results and interpretation.** Table 8 shows the output for a single multiple regression model with the same data as in Table 7 but this time using dummy coding, so that three first-order effects, three second-order effects, and one third-order effect are estimated.

**Table 8. Annotated regression output for dummy-coded variables**

The REG Procedure

Model: MODEL1
Dependent Variable: Y

Number of Observations Read 32

Number of Observations Used 32

Analysis of Variance

| Source | DF | Sum of Squares | Mean Square | F Value | Pr>F |
|---|---|---|---|---|---|
| Model | 7 | 1686.53173 | 240.93310 | 217.32 | <.0001 |
| Error | 24 | 26.60724 | 1.10863 | | |
| Corrected Total | 31 | 1713.13897 | | | |

| | |
|---|---|
| Root MSE | 1.05292 |
| R-Square | 0.9845 |
| Dependent Mean | -0.17256 |
| Adj R-Sq | 0.9799 |
| Coeff Var | -610.15872 |

| Variable | DF | Parameter Estimate | Standard Error | t Value | Pr > \|t\| |
|---|---|---|---|---|---|
| Intercept | 1 | -3.62490 | 0.52646 | -6.89 | <.0001 |
| $Z_{NP}$ | 1 | 6.46140 | 0.74452 | 8.68 | <.0001 |
| $Z_{GUM}$ | 1 | -4.88693 | 0.74452 | -6.56 | <.0001 |
| $Z_{CS}$ | 1 | -0.08592 | 0.74452 | -0.12 | 0.9091 |
| $Z_{NPXGUM}$ | 1 | 11.50880 | 1.05292 | 10.93 | <.0001 |
| $Z_{NPXCS}$ | 1 | 0.58648 | 1.05292 | 0.56 | 0.5827 |
| $Z_{GUMXCS}$ | 1 | -2.30917 | 1.05292 | -2.19 | 0.0382 |
| $Z_{NPXGUMXCS}$ | 1 | 2.09224 | 1.48905 | 1.41 | 0.1728 |

Compared to the effect-coded model found in Table 7, the overall $F$ statistic is exactly the same ($F$=217.32, $p$<0.01). In addition, the test for significance for the highest-order effect, $Z_{NP\times GUM\times CS}$, is the same ($p$=0.17). However, this is where the similarities end. All of the other regression weights, standard errors, and hypothesis tests are different and must be interpreted differently. For example, the regression coefficient for $Z_{NP}$ is significant ($t$=8.68, $p$<0.01). This indicates that, on average, the use of a Nicotine Patch increases self-efficacy by 6.46 units compared to no use of a Nicotine Patch, *when Gum and Counseling variables are set to zero.* Stated another way, the $t$ test for this regression coefficient tests the null hypothesis that the effect of Nicotine Patch on self-efficacy *among those who did not receive Gum or Counseling* is zero. This is very different from the classical definition of a main effect and very different from the hypothesis test we are most interested in, as stated above. Note that the regression coefficient $Z_{GUM}$ is statistically significant but its counterpart in effect coding, $X_{GUM}$, was not significant. This illustrates how the results of hypothesis tests of individual effects can be different between effect and dummy coding.

## Regression Software Packages

**SAS (v 9.2).** The analyses reported in the numerical example above were obtained using SAS PROC REG. We recommend this procedure when using SAS because in PROC REG the user is responsible for coding the independent variables. We find with this approach the interpretation of effects is more straightforward.

Another procedure, PROC GLM, is often used to analyze factorial experimental data. In the GLM procedure, the user has the option to use or not to use the CLASS statement. If the CLASS statement is NOT used, the independent variables are considered quantitative (as they are in PROC REG), and the way in which they are coded by the user matters; in other words, different results will be obtained depending on whether effect coding or dummy coding is used. However, if the CLASS statement is used and the user requests a solution for the regression parameter estimates, the output can be confusing. The ANOVA table corresponds to effect coding. However, for the regression parameter estimates, *the SAS default is to report results for dummy coded variables.* Thus the ANOVA table reports different hypothesis tests than those in the regression section of the output. Some readers may be familiar with using the partitioning of the sums of squares to test whether or not an effect in a regression model is significant. In PROC REG, Type I (hierarchical) and Type II (partial) can be requested as additional output. In PROC GLM, Type I (hierarchical) and Type III (partial) are the default; Type II (unique) can be requested as additional output. Table 9 below explains how the coding scheme and use of a CLASS statement can change the meaning of the effects for the PROC GLM procedure. In general, if the PROC GLM procedure is used, we recommend *not* using the CLASS statement and coding the independent variables yourself to avoid confusion.

**Table 9. Interpretations of results from different approaches to conducting ANOVA using PROC GLM**

| Coding scheme | Interpretation of effect estimates | |
|---|---|---|
| Effect coding | Without CLASS statement | With CLASS statement |
| Type I sums of squares | Classical[1] | Classical |
| Type III sums of squares | Classical | Classical |
| *b* weights & Hypothesis tests | Classical | Dummy |
| Dummy coding | | |
| Type I sums of squares | Classical | Classical |
| Type III sums of squares | Dummy[2] | Classical |
| *b* weights & Hypothesis tests | Dummy | Dummy |

[1] Interpreted as classical effects.
[2] Interpreted as first-order, second-order, third-order effects associated with dummy-coded independent variables.

**SPSS (PASW Statistics 18).** As in SAS, there are two options available to users of SPSS to analyze factorial experiments. Users can either invoke the “LINEAR” procedure under the “Regression” pull-down tab or the “UNIVARIATE” procedure under the “General Linear Model” pull-down tab. Similar to SAS PROC REG, the LINEAR procedure assumes the independent variables are quantitative; therefore, how the variables are coded determines whether or not the parameter estimates and corresponding tests reflect the classical definitions. Within the “UNIVARIATE” procedure, the user can enter the variables as covariates or factors. If the variables are entered as covariates, then the way in which the variables are coded matters; it is only with effect coding that the results will be consistent with the classical definitions of effects (see Table 10). However, if the variables are entered as factors, then SPSS will dummy code the variables. The partitioning of the sums of squares is also included in the output for the UNIVARIATE procedure. Type III SS (partial) is the default, although Type I and Type II can be requested. Table 10 demonstrates how the coding scheme and the choice of how the variables are characterized changes the meaning of the effects. To avoid confusion, we recommend that if the UNIVARIATE procedure is used, the independent variables should be entered as covariates, not as factors.

**Table 10. Interpretations of results from different approaches to conducting ANOVA using SPSS UNIVARIATE procedure**

| Coding scheme | Interpretation of effect estimates | |
|---|---|---|
| Effect coding | Variables as covariates | Variables as factors |
| Type III sums of squares | Classical[1] | Classical |
| *b* weights & Hypothesis tests | Classical | Dummy |
| Dummy coding | | |
| Type III sums of squares | Dummy[2] | Classical |
| *b* weights & Hypothesis tests | Dummy | Dummy |

[1] Interpreted as classical effects.
[2] Interpreted as first-order, second-order, third-order effects associated with dummy-coded independent variables.

## Discussion

In this technical report we have presented a hypothetical example of a $2^3$ factorial experiment to demonstrate that how a categorical variable is coded determines the interpretation of its effects in a regression analysis. Throughout this report, we stressed that differences emerge when product terms are included in the model. So what happens when product terms are not included in the model, that is, when it is assumed that there are no interactions? In this instance, it does not matter what type of coding scheme is used: the test for significance will be the same for each of the effects. However, there will be differences in the scale of the regression coefficients and the interpretation of the intercept term. For the dummy-coded model, the regression coefficient represents the classically defined main effects, whereas with effect coding, the regression coefficient needs to be multiplied by a scaling factor of 2 in order to be equivalent to the classically defined main effect.

In the beginning of this technical report we noted that the report is based on experiments with balanced data. This was done to avoid making the discussion overly complex. The general principles discussed here hold even when the number of subjects is not the same across experimental conditions.

## Take-Away Messages

The following are some take-away messages of this technical report.

1) Different coding schemes for categorical variables estimate different effects, and therefore can result in different results and different conclusions.
2) If dummy coding is used and product terms are included in an analysis, the interpretations of the regression coefficients do not correspond to the classical definitions of main effects and interaction effects. Rather, the effects in a regression with dummy-coded variables should be interpreted as the effect of the variable, when all other variables in the analysis are set to zero (i.e., a simple effect); the only exception is the highest order effect.
3) If effect coding is used and product terms are modeled in an analysis, the interpretations of the regression coefficients *do* correspond to the classical definitions of main effects and interaction effects.
4) Because of this difference in interpretation, we encourage researchers to be very specific about what kind of coding they use, and to be precise about the meaning of the effects they are modeling.
5) There are numerous statistical software options available to researchers to analyze data from factorial experiments using multiple regression. Researchers should be aware of the default options and how use of these defaults may affect their results.

## References

Baker, T. B., Mermelstein, R. J., Collins, L. M. , Piper, M. E., Jorenby, D. E., Smith, S. S., … Fiore, M. C. (2011). New methods for tobacco dependence treatment research. *Annals of Behavioral Medicine, 41*(2), 192-207.

Chakraborty, B., Collins, L. M., Strecher, V. J., & Murphy, S. A. (2009). Developing multicomponent interventions using fractional factorial designs. *Statistics in Medicine, 28,* 2687-2708.

Collins, L. M., Baker, T. B., Mermelstein, R. J., Piper, M. E., Jorenby, D. E., Smith, S. S., … Fiore, M. C. (2011). The multiphase optimization strategy for engineering effective tobacco use interventions. *Annals of Behavioral Medicine, 41*(2), 208-226.

Fisher, R. A. (1925). *Statistical methods for researchers.* Edinburgh, Scotland: Oliver and Boyd.

Fisher, R. A. (1971). *The design of experiments*. New York, NY: Hafner Publishing.

Hardy, M. A. (1993). *Regression with dummy variables.* Newbury Park, CA: Sage.

Montgomery, D. C. (2009). *Design and analysis of experiments*. Hoboken, NJ: John Wiley & Sons, Inc.

Rodgers, J. L., Nicewander, W. A., & Toothaker, L. (1984). Linearly independent, orthogonal, and uncorrelated variables. *The American Statistician, 38*(2), 133-134.

## Acknowledgments

The preparation of this report was supported by National Cancer Institute grant P50 CA143188, National Institute on Drug Abuse grant P50 DA010075-15, and National Institutes of Health grant 8KL2TR000126-02. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Cancer Institute, the National Institute on Drug Abuse, or the National Institutes of Health.

We would like to thank Inbal Nahum-Shani for her suggestion of referring to the effects in a regression with dummy-coded variables as first-order effects, second-order effects, etc. to distinguish them from classically defined main effects and interactions.

**Appendix A. Glossary of Definitions and Terms for $2^k$ Factorial Experiments**

Balanced experiment means that an equal number of subjects are assigned to each study condition.

Main effect, according to the classical definition, is the difference between the mean response at one level of a particular factor and the mean response at the second level, collapsing across the levels of all other factors.

Interaction effect, according to the classical definition, is the average of the difference in the effect of a particular factor across differing levels of a second factor, collapsing over all other factors.

First-order effect (when dummy coding is used), is the difference between the mean response at one level of a particular factor and the mean response at the second level, when all the other factors are set to zero.

Second-order effect (when dummy coding is used), is the average of the difference in the effect of a particular factor across differing levels of a second factor, when all the other factors are set to zero.

Third-order effect (when dummy coding is used), is the average of the difference in the two-way interaction effects at differing levels of the third factor, when all the other factors are set to zero.

## Appendix B. Basis for interpretation of regression coefficients when dummy coding is used

The equations below show the correspondence between the population means and the regression coefficients for dummy-coded variables.

1) Interpretation of $\alpha_1$

$\mu_{(---)} = \alpha_0$: Mean response when A = low, B = low, and C = low

$\mu_{(+--)} = \alpha_0 + \alpha_1$: Mean response when A = high, B = low, and C = low

Therefore, $\alpha_1 = \mu_{(+--)} - \mu_{(---)}$, that is, the effect of A = high vs. A = low, when B = low and C = low.

Note the similarities and differences between this and the classical definition of a main effect of A: $\bar{\mu}_{(+\cdot\cdot)} - \bar{\mu}_{(-\cdot\cdot)}$.

2) Interpretation of $\alpha_{12}$

$\mu_{(---)} = \alpha_0$: Mean response when A = low, B = low, and C = low

$\mu_{(+--)} = \alpha_0 + \alpha_1$: Mean response when A = high, B = low, and C = low

$\mu_{(-+-)} = \alpha_0 + \alpha_2$: Mean response when A = low, B = high, and C = low

$\mu_{(+--)} = \alpha_0 + \alpha_1 + \alpha_2 + \alpha_{12}$: Mean response when A = high, B = high, and C = low

Therefore, $\alpha_{12} = \mu_{(++-)} - \mu_{(+--)} - \mu_{(-+-)} + \mu_{(---)} = \left(\mu_{(++-)} - \mu_{(-+-)}\right) - \left(\mu_{(+--)} - \mu_{(---)}\right)$, that is, the difference in the effect of A = high vs. A = low when B = high, C = low and B = low, C = low.

Note the similarities and differences between this and the classical definition of a two-way interaction effect of A and B: $A \times B = \frac{1}{2}\left[\left(\bar{\mu}_{(++\cdot)} - \bar{\mu}_{(-+\cdot)}\right) - \left(\bar{\mu}_{(+-\cdot)} - \bar{\mu}_{(--\cdot)}\right)\right]$.

3) Interpretation of $\alpha_{123}$

$\mu_{(---)} = \alpha_0$: Mean response when A = low, B = low, and C = low

$\mu_{(+--)} = \alpha_0 + \alpha_1$: Mean response when A = high, B = low, and C = low

$\mu_{(-+-)} = \alpha_0 + \alpha_2$: Mean response when A = low, B = high, and C = low

$\mu_{(++-)} = \alpha_0 + \alpha_1 + \alpha_2 + \alpha_{12}$: Mean response when A = high, B = high, and C = low

$\mu_{(--+)} = \alpha_0 + \alpha_3$: Mean response when A = low, B = low, and C = high.

$\mu_{(+-+)} = \alpha_0 + \alpha_1 + \alpha_3 + \alpha_{13}$: Mean response when A = high, B = low, and C = high

$\mu_{(-++)} = \alpha_0 + \alpha_2 + \alpha_3 + \alpha_{23}$: Mean response when A = low, B = high, and C = high

$\mu_{(+++)} = \alpha_0 + \alpha_1 + \alpha_2 + \alpha_3 + \alpha_{12} + \alpha_{13} + \alpha_{23} + \alpha_{123}$: Mean response when A = high, B = high, and C = high

Therefore,

$$\alpha_{123} = \mu_{(+++)} - \mu_{(-++)} - \mu_{(+-+)} - \mu_{(++-)} + \mu_{(+--)} + \mu_{(-+-)} + \mu_{(--+)} - \mu_{(---)} = \\ \left[\left(\mu_{(+++)} - \mu_{(-++)}\right) - \left(\mu_{(+-+)} - \mu_{(--+)}\right)\right] - \left[\left(\mu_{(++-)} - \mu_{(-+-)}\right) - \left(\mu_{(+--)} - \mu_{(---)}\right)\right].$$

That is, the difference in the interaction effect of A and B when C = high and C = low.

Note the similarities and differences between this and the classical definition of a three-way interaction effect of A, B, and C:

$$A \times B \times C \; = \frac{1}{4}\left[\left[\left(\mu_{(+++)} - \mu_{(-++)}\right) - \left(\mu_{(+-+)} - \mu_{(--+)}\right)\right] - \left[\left(\mu_{(++-)} - \mu_{(-+-)}\right) - \left(\mu_{(+--)} - \mu_{(---)}\right)\right]\right].$$

4) Another way to look at $\alpha_1$ (as shown on Page 14)

$$\alpha_1 = \text{ME}_A - \left(\frac{1}{2}\alpha_{12} + \frac{1}{2}\alpha_{13} + \frac{1}{4}\alpha_{123}\right) \; (*)$$

$$\text{ME}_A = \bar{\mu}_{(+\cdot\cdot)} - \bar{\mu}_{(-\cdot\cdot)}$$

$$\alpha_1 = \mu_{(+--)} - \mu_{(---)} \quad (**)$$

$$\alpha_{12} = \mu_{(++-)} - \mu_{(-+-)} - \mu_{(+--)} + \mu_{(---)} = \left(\mu_{(++-)} - \mu_{(-+-)}\right) - \left(\mu_{(+--)} - \mu_{(---)}\right)$$

$$\alpha_{12} = \mu_{(+-+)} - \mu_{(--+)} - \mu_{(+--)} + \mu_{(---)} = \left(\mu_{(+-+)} - \mu_{(--+)}\right) - \left(\mu_{(+--)} - \mu_{(---)}\right)$$

$$\begin{aligned}\alpha_{123} &= \mu_{(+++)} - \mu_{(-++)} - \mu_{(+-+)} - \mu_{(++-)} + \mu_{(+--)} + \mu_{(-+-)} + \mu_{(--+)} - \mu_{(---)} \\ &= \left[\left(\mu_{(+++)} - \mu_{(-++)}\right) - \left(\mu_{(+-+)} - \mu_{(--+)}\right)\right] - \left[\left(\mu_{(++-)} - \mu_{(-+-)}\right) - \left(\mu_{(+--)} - \mu_{(---)}\right)\right]\end{aligned}$$

So, $(*)$ becomes

$$\begin{aligned}\alpha_1 = \frac{1}{4}&\left\{\left[\mu_{(+++)} + \mu_{(++-)} + \mu_{(+-+)} + \mu_{(+--)}\right] - \left[\mu_{(-++)} + \mu_{(-+-)} + \mu_{(--+)} + \mu_{(---)}\right]\right\} \\ -\frac{1}{2}&\left\{\left[\mu_{(++-)} - \mu_{(-+-)}\right] - \left[\mu_{(+--)} - \mu_{(---)}\right]\right\} - \frac{1}{2}\left\{\left[\mu_{(+-+)} - \mu_{(--+)}\right] - \left[\mu_{(+--)} - \mu_{(---)}\right]\right\} \\ -\frac{1}{4}&\left\{\left[\left(\mu_{(+++)} - \mu_{(-++)}\right) - \left(\mu_{(+-+)} - \mu_{(--+)}\right)\right] - \left[\left(\mu_{(++-)} - \mu_{(-+-)}\right) - \left(\mu_{(+--)} - \mu_{(---)}\right)\right]\right\}\end{aligned}$$

$$\begin{aligned}&= \frac{1}{4}\mu_{(+++)} + \frac{1}{4}\mu_{(++-)} + \frac{1}{4}\mu_{(+-+)} + \frac{1}{4}\mu_{(+--)} - \frac{1}{4}\mu_{(-++)} - \frac{1}{4}\mu_{(-+-)} - \frac{1}{4}\mu_{(--+)} - \frac{1}{4}\mu_{(---)} - \frac{1}{2}\mu_{(++-)} \\ &\quad + \frac{1}{2}\mu_{(-+-)} + \frac{1}{2}\mu_{(+--)} - \frac{1}{2}\mu_{(---)} - \frac{1}{2}\mu_{(+-+)} + \frac{1}{2}\mu_{(--+)} + \frac{1}{2}\mu_{(+--)} - \frac{1}{2}\mu_{(---)} - \frac{1}{4}\mu_{(+++)} + \frac{1}{4}\mu_{(-++)} \\ &\quad + \frac{1}{4}\mu_{(+-+)} - \frac{1}{4}\mu_{(--+)} + \frac{1}{4}\mu_{(++-)} - \frac{1}{4}\mu_{(-+-)} - \frac{1}{4}\mu_{(+--)} + \frac{1}{4}\mu_{(---)}\end{aligned}$$

$$= \mu_{(+--)} - \mu_{(---)},$$

which is exactly the same equation in $(**)$.

Note: In the above equation, the colors denote items from the same experimental condition.