SPSS and R Illustrations for Creating Dummy and Effect Codes

SPSS Syntax

I suggest meaningful codes for dummies, so you can easily match the results to the data when you view the output. It is best to name each contrast dummy variable after the specific comparison (e.g., SVQ = social versus quant) or just after the group (e.g., SOCIAL, assuming you know the referent is the quant group). The variables starting with E are unweighted effect codes, and those starting with W are weighted effect codes.

```
do if area eq 1.
compute SVQ=1.
compute DVQ=0.
compute IVQ=0.
compute E1=1.
compute E2=0.
compute E3=0.
compute W1=1.
compute W2=0.
compute W3=0.
else if area eq 2.
compute SVQ=0.
compute DVQ=1.
compute IVQ=0.
compute E1=0.
compute E2=1.
compute E3=0.
compute W1=0.
compute W2=1.
compute W3=0.
else if area eq 3.
compute SVQ=0.
compute DVQ=0.
compute IVQ=1.
compute E1=0.
compute E2=0.
compute E3=1.
compute W1=0.
compute W2=0.
compute W3=1.
else if area eq 4.
compute SVQ=0.
compute DVQ=0.
compute IVQ=0.
compute E1=-1.
compute E2=-1.
compute E3=-1.
compute W1=-1.5.
compute W2=-1.25.
compute W3=-1.25.
end if.
```

R Code Example (Dummy Codes Only)

I always encourage students to create their own dummy variables. R will create them for you automatically if your multiple-category variable is a "factor" (nominal variable) in the data set. Here is what I get with the psychology area and satisfaction example.

```
library(lessR)
reg.brief(satis ~ area)

BASIC ANALYSIS

Estimated Model
```

	Estimate	Std Err	t-value	p-value	Lower 95%	Upper 95%
(Intercept)	82.667	3.254	25.407	0.000	75.769	89.564
areaDevelopmental	-1.067	4.826	-0.221	0.828	-11.297	9.164
areaI/O	-1.467	4.826	-0.304	0.765	-11.697	8.764
areaQuant	-3.167	5.145	-0.616	0.547	-14.073	7.739

Messy, huh? It picks my referent for me and it reverses the signs of all of the coefficients.
Below, I create my own dummy variables. I like my variables numeric, so I transform them first (using lessR). The ifelse function states that if the area variable is equal to sum number (e.g., mydata$area==1 in the first line), then code my new variable (svq) as 1, else code the new variable as 0.

```
.
library(lessR)
mydata <-Transform(area = as.numeric(area))

mydata$svq<-ifelse(mydata$area==1, 1,0)
mydata$dvq<-ifelse(mydata$area==2, 1,0)
mydata$ivq<-ifelse(mydata$area==3, 1,0)
mydata
```

```
R-squared:   0.024     Adjusted R-squared:   -0.160     PRESS R-squared:   -0.599

Null hypothesis that all population slope coefficients are 0:
  F-statistic: 0.128     df: 3 and 16     p-value:  0.942
```

Analysis of Variance

	df	Sum Sq	Mean Sq	F-value	p-value
svq	1	13.752	13.752	0.217	0.648
dvq	1	4.292	4.292	0.068	0.798
ivq	1	6.422	6.422	0.101	0.755
Model	3	24.467	8.156	0.128	0.942