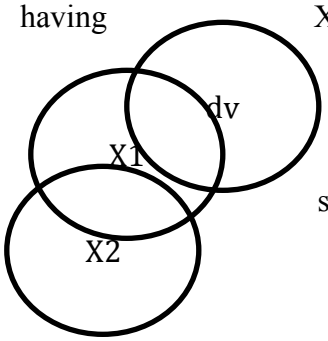
Handout R13A
Psyx 501 Prof. Colleen F Moore
Montana State University (University of Wisconsin)
Spring 2014

An Oversimplified Example of Suppression
(drawn partly from Cohen & Cohen, 1975, pp. 84 ff, and Cohen & Cohen, 1983, pp.95 ff)

I. The term 'suppression' refers to situations in which the relationships among the IVs hide (suppress) their actual relationships to the DV. Cohen and Cohen (1975) draw these diagrams to illustrate (Fig. 3.4.1, p. 86).

Here the dv is predictable from X1, but not X2. Cohen and Cohen call this 'classical suppression'.
The correlation matrix will show a correlation between the dv and X1, and a correlation between the X1 and X2, but no correlation between X2 and dv. In the multiple regression having X2 in the model will actually increase the variance in the dv that is explained. Cohen and Cohen say this is because X2 is 'suppressing' the variance in X1 that is not relevant to predicting the dv. They give a numerical example using the semi-partial correlations. They show that the sign of the semipartial correlation of X2 with dv actually flips direction!



II. Example calculated in R.

I took the data from Cohen & Cohen, 2nd ed, p. 82 and changed it to illustrate suppression. These are faculty salaries (from the 1980s), and we wish to predict salary from years since PhD and number of publications.

```
> cohen3=read.table(pipe("pbpaste"),header=T)
> cohen3
   salary numpub yrsphd
1   18000      2     10
2   19961      4      2
3   19828      5      5
4   17030     12     10
5   19925      5     10
6   19041      9      7
7   27132      3      4
8   27268      1      2
9   32483      8      6
10  27029     12      8
11  25362      9      9
12  28463      4      8
```

```
13  32931      8      8
14  28270     11      7
15  38362     21     10

> summary(cohen3)
     salary          numpub         yrsphd
 Min.   :17030   Min.   : 1.0   Min.   : 2.000
 1st Qu.:19876   1st Qu.: 4.0   1st Qu.: 5.500
 Median :27029   Median : 8.0   Median : 8.000
 Mean   :25406   Mean   : 7.6   Mean   : 7.067
 3rd Qu.:28366   3rd Qu.:10.0   3rd Qu.: 9.500
 Max.   :38362   Max.   :21.0   Max.   :10.000

> attach(cohen3)
> cor(cohen3)
           salary    numpub     yrsphd
salary 1.00000000 0.4606314 0.01643274
numpub 0.46063141 1.0000000 0.51508324
yrsphd 0.01643274 0.5150832 1.00000000

# notice that the correlation between yrsphd and salary is
zero. The bivariate correlations are like the diagram above

# first fit simple regression with numpub as IV
> mod2=lm(salary~numpub); summary(mod2)

Call:
lm(formula = salary ~ numpub)

Residuals:
     Min       1Q   Median       3Q      Max
-10864.9  -4058.2   -835.7   5234.7   7299.0

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  21106.0     2744.5   7.690 3.44e-06 ***
numpub         565.7      302.3   1.871    0.084 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’
1

Residual standard error: 5813 on 13 degrees of freedom
Multiple R-squared:  0.2122,  Adjusted R-squared:  0.1516
F-statistic: 3.501 on 1 and 13 DF,  p-value: 0.084

# The R-squared is .21, and the effect of numpub misses
significance.
```

Note: If you have missing data you need to add a ‘use=’ option to the cor statement.

```
# Now add yrsphd to the regression equation.
> mod3=lm(salary~numpub+yrsphd)
> summary(mod3)

Call:
lm(formula = salary ~ numpub + yrsphd)

Residuals:
   Min     1Q Median     3Q    Max
 -9688  -3266    225   4107   7864

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  24510.9     4239.5   5.782 8.72e-05 ***
numpub         755.9      351.3   2.151   0.0525 .
yrsphd        -686.3      653.2  -1.051   0.3141
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’
1

Residual standard error: 5789 on 12 degrees of freedom
Multiple R-squared:  0.2786,  Adjusted R-squared:  0.1583
F-statistic: 2.317 on 2 and 12 DF,  p-value: 0.141

## whoa! Look at the coefficient on yrsphd: -686.3. This
says that holding numpub constant, each year since PhD
reduces salary by $686. Also, the coefficient on numpub is
now larger. Each publication predicts an increase in salary
of about $756. Without yrsphd in the model, each
publication predicts an increase of $566 in the model.
```

III. Further remarks.

Cohen & Cohen (2nd edition, p. 96) comment on possible causation: “Suppression is a plausible model for many homeostatic mechanisms, both biological and social, in which force and counterforce tend to occur together and have counteractive effects.” They go on to talk about the timing of the measurements in such systems and finding equilibrium points.

My take-away message from this section is that there is **no escaping experimental design**, and the timing of measurements is a key issue.