Example models (abridged)

Thanks to Kara Weisman (2014)

Setup

These are real pilot data, collected on 2014-10-18, exploring people's implicit theories of selfhood, personhood, personal identity, etc. - "Experimental Philosophy"!

Here, we'll focus on trying to predict a subject's comfort with the idea of neuron replacement surgery, in which individual neurons are replaced with "electronic circuits" that function just like neurons. The "comfort" score is a composite score, including responses to questions like "If your doctor recommended this procedure (e.g., to prevent dementia), would you have it done?".

Variables:

- `Subj`: unique identifier for subject
- `Cond`: 2 versions of teleporter scenario, C1 (coded as 0) vs. C2 (coded as 1)
- `Gender`: male or female
- `Relig`: Judeo-Christian (coded as 1) or not religious (coded as 0)
- `EnjoyScifi`: how much subj enjoys scifi (range: [-4, +4])
- `ThinkOften`: how often subj thinks about these thought expts (range: [0, 6])
- `Teleporter`: comfort with teleporter scenario (range: [-15, +15])
- `NeuronRepl`: comfort with the idea of neuron replacement surgery scenario (range: [-12, +12])

```r
# Load in packages
library(ggplot2)
library(scatterplot3d)

# Load in data
xphil = read.csv("http://web.stanford.edu/~kweisman/XPhil/xphil_demo.csv")

# Remove Relig NAs
xphil = subset(xphil, Religion2 != "NA")

# Recode variables to have shorter names
xphil$Cond = factor(xphil$Condition, labels = c("C1", "C2"))
xphil$Relig = factor(xphil$Religion2, labels = c("JC", "None"))

# Set contrasts
contrasts(xphil$Relig) = c(1,0) # changed
```

Single regression (1 predictor)

One continuous (numeric) predictor

Research question: Does a subj's comfort with the idea of neuron replacement surgery depend on how much s/he enjoys science fiction?

```
# Effect of EnjoyScifi on NeuronRepl
r1 = lm(NeuronRepl ~ EnjoyScifi, data = xphil); summary(r1)

##
## Call:
## lm(formula = NeuronRepl ~ EnjoyScifi, data = xphil)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -13.35  -3.75   1.60   4.30   9.20
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)    3.750      0.775    4.84  7.2e-06 ***
## EnjoyScifi     0.650      0.284    2.29    0.025 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.56 on 73 degrees of freedom
## Multiple R-squared:  0.0672, Adjusted R-squared:  0.0544
## F-statistic: 5.26 on 1 and 73 DF,  p-value: 0.0247
```

What is the regression equation?

What does the intercept mean?

What does the coefficient for `EnjoyScifi` *mean?*

How would you write up these results?

One discrete (categorical) predictor

Research question: Does a subj's comfort with the idea of neuron replacement surgery depend on his/her gender?

```
# Effect of Gender on NeuronRepl
r2a = lm(NeuronRepl ~ Gender, data = xphil); summary(r2a)

##
## Call:
## lm(formula = NeuronRepl ~ Gender, data = xphil)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -12.63  -4.06   1.37   4.87   7.44
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)     3.56       1.00    3.56  0.00066 ***
## GenderM         2.07       1.32    1.56  0.12263
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.66 on 73 degrees of freedom
## Multiple R-squared:  0.0323, Adjusted R-squared:  0.0191
## F-statistic: 2.44 on 1 and 73 DF,  p-value: 0.123
```

What is the regression equation?

What does the intercept mean?

What does the coefficient for `GenderM` *mean?*

What is the mean comfort level for men? For women?

How would you write up these results?

...what about his/her religion?

```
# Effect of Relig on NeuronRepl
r2b = lm(NeuronRepl ~ Relig, data = xphil); summary(r2b)

##
## Call:
## lm(formula = NeuronRepl ~ Relig, data = xphil)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -12.39  -4.74   1.60   4.92   7.92
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)    5.395      0.928    5.81  1.5e-07 ***
## Relig1        -1.314      1.321   -0.99     0.32
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.72 on 73 degrees of freedom
## Multiple R-squared:  0.0134, Adjusted R-squared:  -0.000148
## F-statistic: 0.989 on 1 and 73 DF,  p-value: 0.323
```

What is the regression equation?

What does the intercept mean?

What does the coefficient for `Relig1` *mean?*

What is the mean comfort level for non-religious people? For Judeo-christian people?

How would you write up these results?

...what about his/her experimental condition?

```
# Effect of Cond on NeuronRepl
r2c = lm(NeuronRepl ~ Cond, data = xphil); summary(r2c)

##
## Call:
## lm(formula = NeuronRepl ~ Cond, data = xphil)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -12.34  -4.74   1.66   4.87   7.87
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)    5.342      0.929    5.75  1.9e-07 ***
## CondC2        -1.207      1.322   -0.91     0.36
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.73 on 73 degrees of freedom
## Multiple R-squared:  0.0113, Adjusted R-squared:  -0.00226
## F-statistic: 0.833 on 1 and 73 DF,  p-value: 0.364
```

What is the regression equation?

What does the intercept mean?

What does the coefficient for `CondC2` *mean?*

What is the mean comfort level for people in Condition 1? For people in Condition 2?

How would you write up these results?

Multiple regression (≥ 1 predictor)

Additive models

Two continuous predictors

Research question: Does a subj's comfort with the idea of neuron replacement surgery depend his/her enjoyment of science fiction, and/or his/her comfort with teleportation (the thought experiment presented immediately prior to this judgment)?

```
# Effects of EnjoyScifi + Teleporter on NeuronRepl
r3a = lm(NeuronRepl ~ EnjoyScifi + Teleporter, data = xphil); summary(r3a)

##
## Call:
## lm(formula = NeuronRepl ~ EnjoyScifi + Teleporter, data = xphil)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -12.78  -3.83   1.16   4.00   8.46
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.6075     0.8519    3.06   0.0031 **
## EnjoyScifi    0.7481     0.2740    2.73   0.0080 **
## Teleporter    0.2038     0.0744    2.74   0.0077 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.33 on 72 degrees of freedom
## Multiple R-squared:  0.155,  Adjusted R-squared:  0.132
## F-statistic: 6.62 on 2 and 72 DF,  p-value: 0.00229
```

What is the regression equation?

What does the intercept mean?

What do the coefficients for `EnjoyScifi` *and* `Teleporter` *mean?*

How would you write up these results?

Two categorical predictors

Research question: Does a subj's comfort with the idea of neuron replacement surgery depend his/her experimental condition, and/or his/her religion?

```
# Effects of Cond + Relig on NeuronRepl
r3b = lm(NeuronRepl ~ Cond + Relig, data = xphil); summary(r3b)

##
## Call:
## lm(formula = NeuronRepl ~ Cond + Relig, data = xphil)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -11.78  -4.73   1.32   4.77   8.62
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)     6.08       1.16    5.24  1.5e-06 ***
## CondC2         -1.30       1.32   -0.98     0.33
## Relig1         -1.40       1.32   -1.06     0.29
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.72 on 72 degrees of freedom
## Multiple R-squared:  0.0264, Adjusted R-squared:  -0.000633
## F-statistic: 0.977 on 2 and 72 DF,  p-value: 0.382
```

What is the regression equation?

What does the intercept mean?

What do the coefficients for `CondC2` *and* `Relig1` *mean?*

What is the mean comfort level for non-religious people in Condition 1? For religious people in Condition 1? For non-religious people in Condition 2? For non-religious people in Condition 2?

How would you write up these results?

One continuous predictor, one categorical predictor

Research question: Does a subj's comfort with the idea of neuron replacement surgery depend his/her religion, and/or his/her enjoyment of scifi?

```
# Effects of Relig + EnjoyScifi on NeuronRepl
r3c = lm(NeuronRepl ~ Relig + EnjoyScifi, data = xphil); summary(r3c)

##
## Call:
## lm(formula = NeuronRepl ~ Relig + EnjoyScifi, data = xphil)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -13.76  -4.04   1.82   4.22   9.56
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)    4.262      1.044    4.08  0.00011 ***
## Relig1        -0.957      1.299   -0.74  0.46380
## EnjoyScifi     0.624      0.287    2.17  0.03294 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.58 on 72 degrees of freedom
## Multiple R-squared:  0.0742, Adjusted R-squared:  0.0485
## F-statistic: 2.88 on 2 and 72 DF,  p-value: 0.0624
```

What is the regression equation?

What does the intercept mean?

What do the coefficients for `Relig1` *and* `EnjoyScifi` *mean?*

How would you write up these results?

Polynomial effects (a kind of "higher-order" term)

Research question: Does the effect of a subject's enjoyment of scifi on his/her comfort with the idea of neuron replacement surgery have a non-linear component?

```
# Effects of poly(EnjoyScifi, 2) on NeuronRepl
r4a = lm(NeuronRepl ~ poly(EnjoyScifi, 2), data = xphil); summary(r4a)

##
## Call:
## lm(formula = NeuronRepl ~ poly(EnjoyScifi, 2), data = xphil)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -14.43  -4.26   1.35   4.20   8.74
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)              4.747      0.636    7.47  1.5e-10 ***
## poly(EnjoyScifi, 2)1    12.753      5.504    2.32    0.023 *
## poly(EnjoyScifi, 2)2     8.733      5.504    1.59    0.117
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.5 on 72 degrees of freedom
## Multiple R-squared:  0.0987, Adjusted R-squared:  0.0737
## F-statistic: 3.94 on 2 and 72 DF,  p-value: 0.0237
```

What does the intercept mean?

What do the coefficients for `poly(EnjoyScif, 2)1` *and* `poly(EnjoyScifi, 2)2` *mean?*

How would you write up these results?

Interactions

Between continuous and categorical predictors

Research question: Does the (linear) effect of a subject's enjoyment of scifi on his/her comfort with the idea of neuron replacement surgery depend on...

...the subject's condition?

```
r5a = lm(NeuronRepl ~ EnjoyScifi * Cond, data = xphil); summary(r5a)

##
## Call:
## lm(formula = NeuronRepl ~ EnjoyScifi * Cond, data = xphil)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -13.43  -3.43   1.55   4.07   8.76
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)          4.807      1.151    4.18  8.3e-05 ***
## EnjoyScifi           0.323      0.429    0.75     0.45
## CondC2              -1.914      1.566   -1.22     0.23
## EnjoyScifi:CondC2    0.561      0.574    0.98     0.33
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.58 on 71 degrees of freedom
## Multiple R-squared:  0.0879, Adjusted R-squared:  0.0494
## F-statistic: 2.28 on 3 and 71 DF,  p-value: 0.0865
```

What does the intercept mean?

What do the coefficients for `EnjoyScifi`, `CondC2`, *and* `EnjoyScifi:CondC2` *mean?*

How would you write up these results?

... the subject's religion?

```
r5b = lm(NeuronRepl ~ EnjoyScifi * Relig, data = xphil); summary(r5b)
```

```
##
## Call:
## lm(formula = NeuronRepl ~ EnjoyScifi * Relig, data = xphil)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -14.87  -3.82   1.26   4.20   7.99
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)          3.340      1.198    2.79   0.0068 **
## EnjoyScifi           1.132      0.437    2.59   0.0117 *
## Relig1               0.427      1.574    0.27   0.7867
## EnjoyScifi:Relig1   -0.879      0.575   -1.53   0.1309
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.53 on 71 degrees of freedom
## Multiple R-squared:  0.104,  Adjusted R-squared:  0.0658
## F-statistic: 2.74 on 3 and 71 DF,  p-value: 0.0498
```

What does the intercept mean?

What do the coefficients for `EnjoyScifi`, `Relig1`, *and* `EnjoyScifi:Relig1` *mean?*

If the last term were significant (e.g., p = 0.001 instead of p = 0.131), which religious group would you conclude is more strongly affected by their enjoyment of science fiction?

How would you write up these results?

What about the *non-linear* effect of the subject's enjoyment of scifi? Does it vary by religion?

```
r5c = lm(NeuronRepl ~ poly(EnjoyScifi,2) * Relig, data = xphil); summary(r5c)
```

```
##
## Call:
## lm(formula = NeuronRepl ~ poly(EnjoyScifi, 2) * Relig, data = xphil)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -14.61  -3.75   1.32   4.23   9.23
##
## Coefficients:
##                             Estimate Std. Error t value Pr(>|t|)
## (Intercept)                    5.053      0.891    5.67    3e-07 ***
## poly(EnjoyScifi, 2)1          22.571      8.516    2.65     0.01 **
## poly(EnjoyScifi, 2)2          -2.270      8.273   -0.27     0.78
## Relig1                        -0.993      1.265   -0.79     0.43
## poly(EnjoyScifi, 2)1:Relig1  -15.914     11.167   -1.43     0.16
## poly(EnjoyScifi, 2)2:Relig1   18.370     11.037    1.66     0.10
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.42 on 69 degrees of freedom
## Multiple R-squared:  0.163,  Adjusted R-squared:  0.103
## F-statistic:  2.7 on 5 and 69 DF,  p-value: 0.0277
```

What does the intercept mean?

What do all the other coefficients mean?

Consider the last term to be marginally significant: What is different about the relationship between enjoyment of science fiction and comfort with the neuron replacement surgery for Judeo-Christian vs. non-religious people?

How would you write up these results?

Among categorical predictors

Research question: Does the effect of a subject's religion on his/her comfort with neuron replacement surgery depend on his/her experimental condition?

```
r6 = lm(NeuronRepl ~ Cond * Relig, data = xphil); summary(r6)

##
## Call:
## lm(formula = NeuronRepl ~ Cond * Relig, data = xphil)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -12.59  -3.99   1.25   4.59   8.35
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)       7.22       1.33    5.42  7.6e-07 ***
## CondC2           -3.47       1.84   -1.89    0.063 .
## Relig1           -3.57       1.84   -1.95    0.056 .
## CondC2:Relig1     4.41       2.62    1.69    0.096 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.65 on 71 degrees of freedom
## Multiple R-squared:  0.0639, Adjusted R-squared:  0.0243
## F-statistic: 1.62 on 3 and 71 DF,  p-value: 0.193
```

What does the intercept mean?

What do all the other coefficients mean?

Consider the last term to be marginally significant: What is different about the effect of condition on comfort with the neuron replacement surgery for Judeo-Christian vs. non-religious people?

How would you write up these results?

Among continuous predictors

Research question: Does the effect of a subject's enjoyment of scifi on his/her comfort with the idea of neuron replacement surgery depend on how often s/he thinks about these kinds of questions?

```
r7 = lm(NeuronRepl ~ EnjoyScifi * ThinkOften, data = xphil); summary(r7)

##
## Call:
## lm(formula = NeuronRepl ~ EnjoyScifi * ThinkOften, data = xphil)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -13.79  -3.85   1.31   4.09   8.31
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)            2.99346    1.15101    2.60    0.011 *
## EnjoyScifi             0.00551    0.42409    0.01    0.990
## ThinkOften             0.10962    0.72421    0.15    0.880
## EnjoyScifi:ThinkOften  0.28688    0.21370    1.34    0.184
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.48 on 71 degrees of freedom
## Multiple R-squared:  0.12,   Adjusted R-squared:  0.0824
## F-statistic: 3.21 on 3 and 71 DF,  p-value: 0.0279

# Note how the "main effects" disappear!
```

What does the intercept mean?

What do all the other coefficients mean?

How would you write up these results?