Moderation Analyses

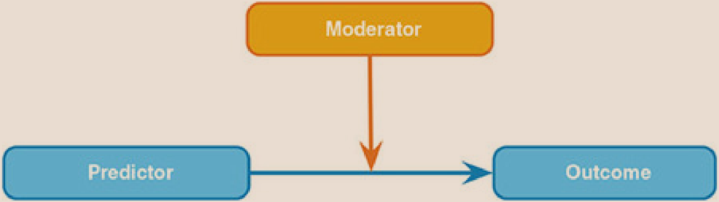
Jonathan A. Pedroza, PhD

Conceptual Model

- conceptually the term is moderation, but its an interaction in statistical terms
  - still the multiplied effect between two IVs
- moderation models allow us to see if there are differences between your predictor and your outcome by a second predictor (**moderator variable**)
  - the moderator affects the relationship between your IV and your DV
    - Ex: you can examine the differences in depression (DV) between BMI categories (IV_1) and whether they differ based on race/ethnicity (IV_2)
    - Ex: you can examine if there are differences by BMI category (IV_2) on the relationship between smartphone use (IV_1) and depression (DV)
    - Ex: you can examine if there are differences in the relationship between smartphone use (IV_1) and depression (DV) by level of email use (IV_2)

Conceptual Model

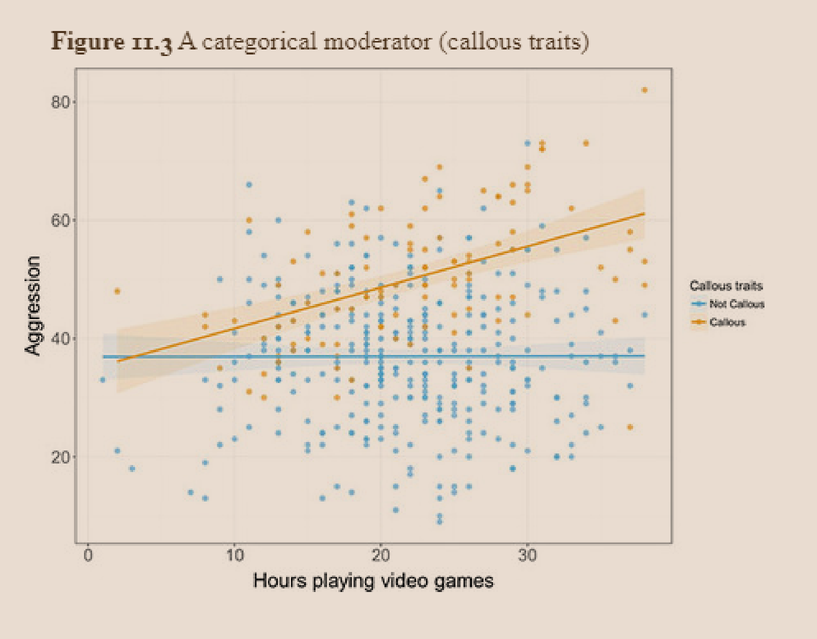
Figure 11.2 Diagram of the *conceptual* moderation model



Conceptual Model

- when we have two continuous variables, we are looking at the relationship between the predictor and your outcome based on chosen levels of your moderator
  - it tends to be +1 and -1 SD for low and high levels of your moderator
- when we look at a continuous predictor and a categorical moderator, we will look at the relationship between the predictor and the outcome compared by groups

Conceptual Model



Conceptual Model

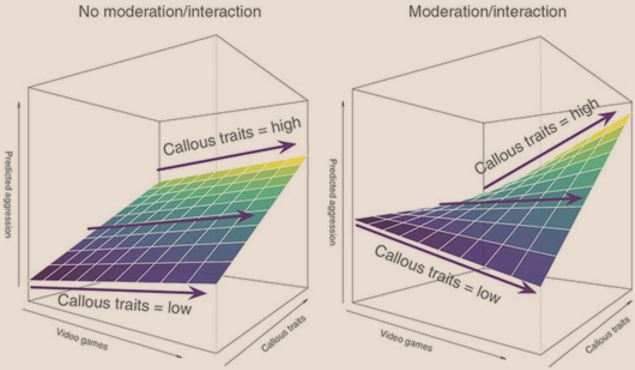


Figure 11.4 A continuous moderator (callous traits)

Statistical Model

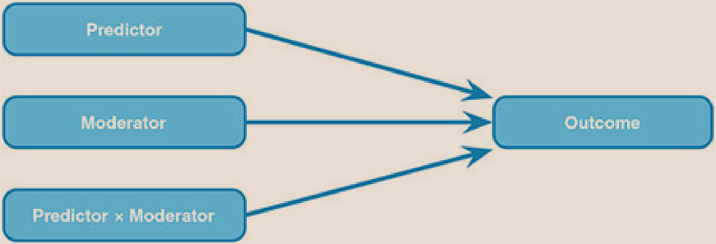


Figure 11.5 Diagram of the *statistical* moderation model

Statistical Model

$$Y_i = (b_0 + b_1 X_{1i} + b_2 X_{2i} + b_3 X_{3i} + b_n X_{ni}) + \epsilon_i$$

$$Stress_i = (b_0 + b_1 Smartphone_i + b_2 Email_i) + \epsilon_i$$

$$Stress_i = (b_0 + b_1 Smartphone_i + b_2 Email_i + b_3 SmartphoneXEmail_i)$$

Centering Variables

- b values represent the slope and should be different from zero in order to reject the null hypothesis
- sometimes the interpretation of your predictors would not make sense at a value of zero
  - Book: you can't have zero heart rate
- Centering is used to redefine the zero point
  - also called **grand mean centering**, because centering is often subtracting the mean from your variable
  - shifts the scale over but retains the unit interpretation
  - the slope won't change between predictor and outcome

Centering Variables

$$Email_{center} = Email - \overline{X}_{Email}$$

- centered values are centered on zero but are slightly different from z-scores
  - centered values are not expressed as standard deviation units (like z-scores)
- centering won't really affect interpretation of the interaction
  - it does affect interpretation of "main effects" relationships between predictors and outcome
  - these relationships are no longer the b values when the other variable(s) are zero
    - they are now the relationship when the other variable(s) are at their mean
- Ex: b_1 would be the relationship between smartphone use and stress when emailing is at its average

Centering Variables

- centering with an interaction makes “main effects” more interpretable, especially with non-significant interactions
- two interpretations are possible for the “main effects”
  - the effect of that predictor at the mean value of the sample
  - the average effect of that predictor across the range of scores for the other predictors

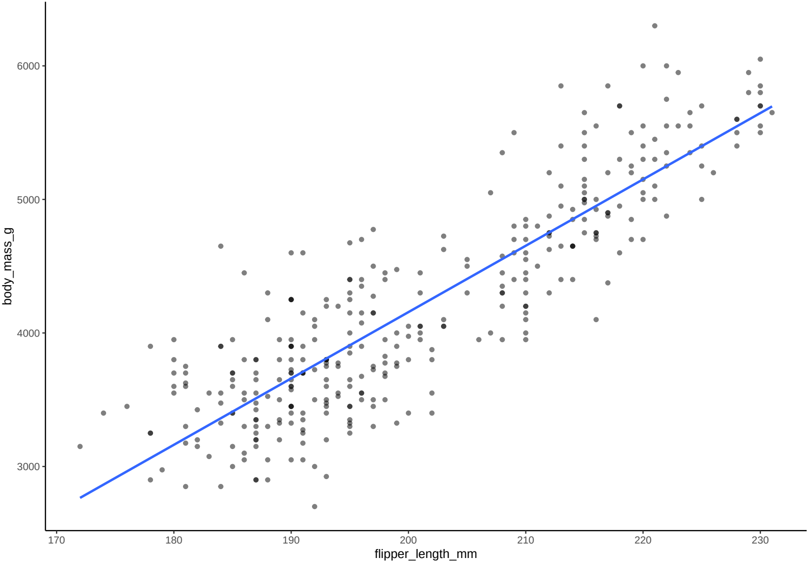
Creating Interaction Variables

- categorical variables will need dummy coding
- continuous variables could be centered
- then you create a term that multiples the two variables of interest
  - for continuous and categorical variable interactions and categorical and categorical interactions –> you will use each dummy coded variable to make interactions

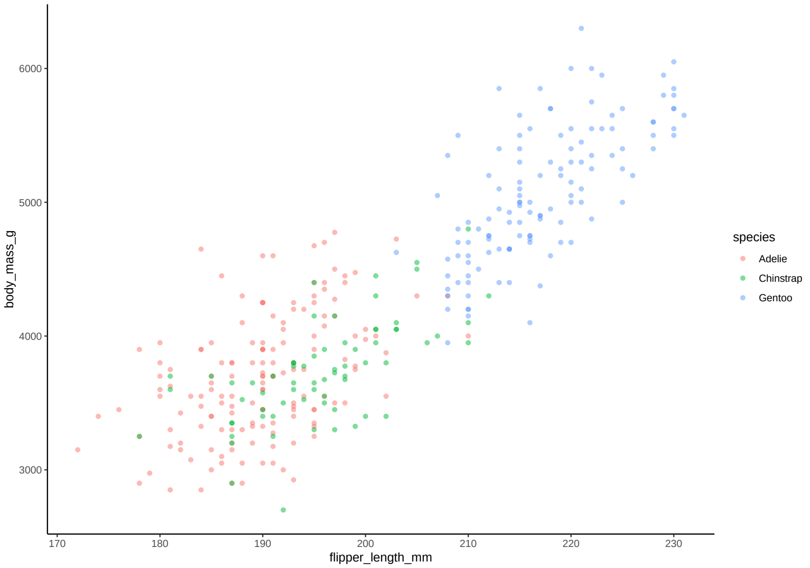
Simple Slopes Analyses

- if you have a significant interaction then you must conduct simple slopes analyses
  - significant interactions only tell us at differing levels of our moderator, there is a significant relationship between predictor and outcome
- we may be interested in the relationship between our predictor and outcome at **high levels** and **low levels** of your moderator
  - this is often at +1 and -1 standard deviations from the centered mean
    - you could use more meaningful values for your moderator, these are just the norm
    - you could also run the interaction at all the varying levels of the moderator rather than just two points
    - **Johnson-Neyman intervals**

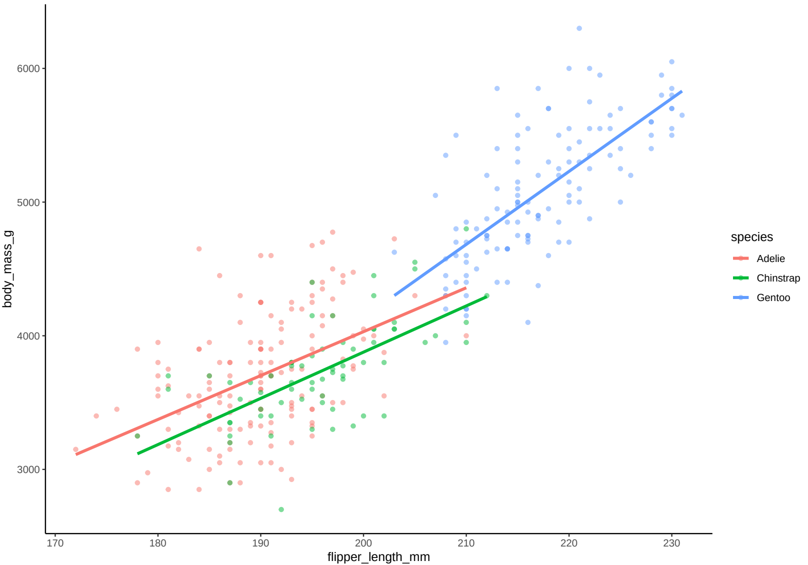
Statistical Model



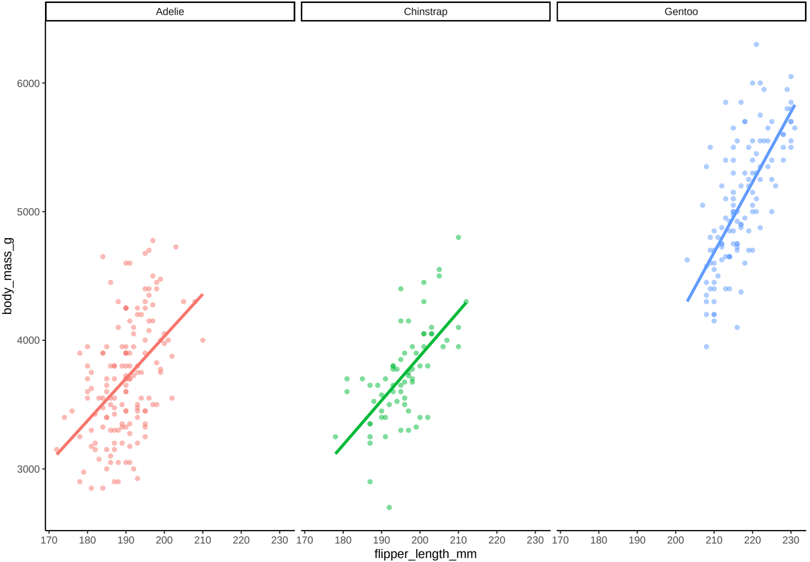
Statistical Model



Statistical Model



Statistical Model



Statistical Model

```
lm(formula = body_mass_g ~ flipper_length_mm, data = penguins)
                  coef.est coef.se  t value  Pr(>|t|)
(Intercept)       -5780.83   305.81   -18.90     0.00
flipper_length_mm    49.69     1.52    32.72     0.00
---
n = 342, k = 2
residual sd = 394.28, R-Squared = 0.76
```

Statistical Model

```
lm(formula = body_mass_g ~ flipper_length_mm + as.factor(species),
    data = penguins)
                             coef.est coef.se  t value  Pr(>|t|)
(Intercept)                  -4031.48   584.15    -6.90     0.00
flipper_length_mm               40.71     3.07    13.25     0.00
as.factor(species)Chinstrap   -206.51    57.73    -3.58     0.00
as.factor(species)Gentoo       266.81    95.26     2.80     0.01
---
n = 342, k = 4
residual sd = 375.54, R-Squared = 0.78
```

Statistical Model

```
lm(formula = body_mass_g ~ flipper_length_mm * as.factor(species),
    data = penguins)
                                             coef.est coef.se  t value
(Intercept)                                  -2535.84  879.47   -2.88
flipper_length_mm                               32.83    4.63    7.10
as.factor(species)Chinstrap                   -501.36 1523.46   -0.33
as.factor(species)Gentoo                     -4251.44 1427.33   -2.98
flipper_length_mm:as.factor(species)Chinstrap    1.74    7.86    0.22
flipper_length_mm:as.factor(species)Gentoo      21.79    6.94    3.14
                                             Pr(>|t|)
(Intercept)                                      0.00
flipper_length_mm                                0.00
as.factor(species)Chinstrap                      0.74
as.factor(species)Gentoo                         0.00
flipper_length_mm:as.factor(species)Chinstrap    0.82
flipper_length_mm:as.factor(species)Gentoo       0.00
---
n = 342, k = 6
residual sd = 370.60, R-Squared = 0.79
```

Statistical Model - Adelie

```
Call:
lm(formula = body_mass_g ~ flipper_length_mm, data = .x)

Residuals:
    Min      1Q  Median      3Q     Max
-875.68 -331.10  -14.53  265.74 1144.81

Coefficients:
                   Estimate Std. Error t value Pr(>|t|)
(Intercept)       -2535.837    964.798  -2.628  0.00948 **
flipper_length_mm    32.832      5.076   6.468 1.34e-09 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 406.6 on 149 degrees of freedom
  (1 observation deleted due to missingness)
Multiple R-squared:  0.2192,    Adjusted R-squared:  0.214
F-statistic: 41.83 on 1 and 149 DF,  p-value: 1.343e-09
```

Statistical Model - Gentoo

```
Call:
lm(formula = body_mass_g ~ flipper_length_mm, data = .x)

Residuals:
    Min      1Q  Median      3Q     Max 
-911.18 -235.76  -51.93  170.75 1015.71 

Coefficients:
                   Estimate Std. Error t value Pr(>|t|)    
(Intercept)       -6787.281   1092.552  -6.212 7.65e-09 ***
flipper_length_mm    54.623      5.028  10.863  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 360.2 on 121 degrees of freedom
  (1 observation deleted due to missingness)
Multiple R-squared:  0.4937,    Adjusted R-squared:  0.4896 
F-statistic:   118 on 1 and 121 DF,  p-value: < 2.2e-16
```

Statistical Model - Chinstrap

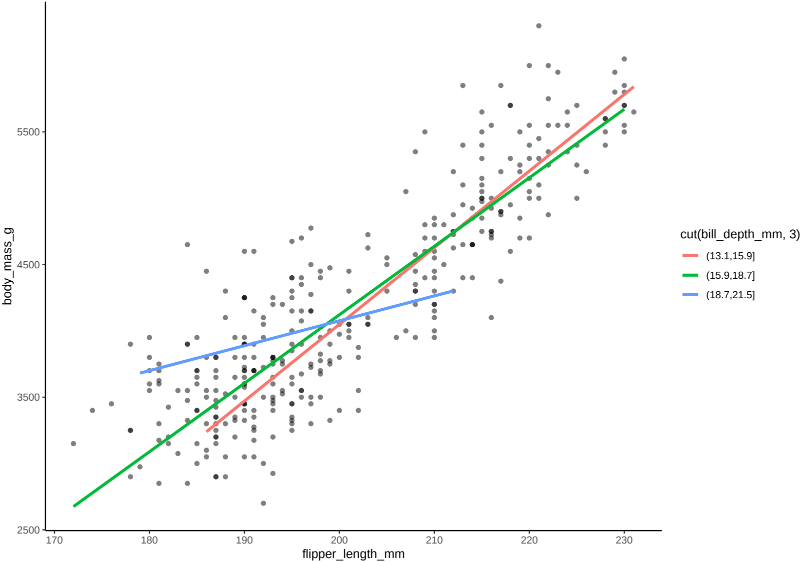
```
Call:
lm(formula = body_mass_g ~ flipper_length_mm, data = .x)

Residuals:
    Min      1Q  Median      3Q     Max 
-900.90 -137.45  -28.55  142.59  695.38 

Coefficients:
                   Estimate Std. Error t value Pr(>|t|)    
(Intercept)       -3037.196    997.054  -3.046  0.00333 ** 
flipper_length_mm    34.573      5.088   6.795 3.75e-09 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 297 on 66 degrees of freedom
Multiple R-squared:  0.4116,	Adjusted R-squared:  0.4027 
F-statistic: 46.17 on 1 and 66 DF,  p-value: 3.748e-09
```

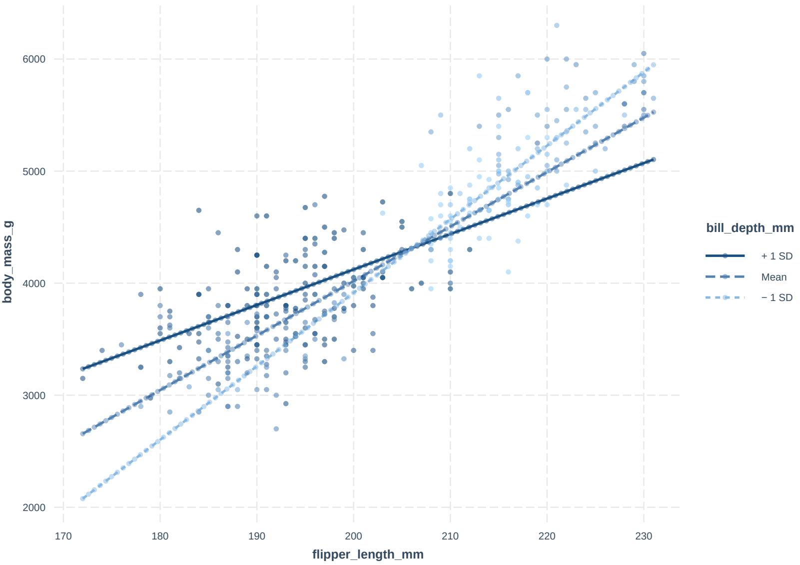
Statistical Model



Statistical Model

```
lm(formula = body_mass_g ~ flipper_length_mm * bill_depth_mm,
    data = penguins)
                                 coef.est  coef.se   t value   Pr(>|t|)
(Intercept)                      -36097.06   4636.27     -7.79      0.00
flipper_length_mm                   196.07     22.60      8.67      0.00
bill_depth_mm                      1771.80    273.00      6.49      0.00
flipper_length_mm:bill_depth_mm      -8.60      1.34     -6.41      0.00
---
n = 342, k = 4
residual sd = 371.78, R-Squared = 0.79
```

Statistical Model



Statistical Model

```
Slope of flipper_length_mm when bill_depth_mm = 15.17638 (- 1 SD):

   Est.    S.E.    t val.       p
------- ------ -------- ------
  65.61   2.81    23.31   0.00

Slope of flipper_length_mm when bill_depth_mm = 17.15117 (Mean):

   Est.    S.E.    t val.       p
------- ------ -------- ------
  48.63   1.82    26.71   0.00

Slope of flipper_length_mm when bill_depth_mm = 19.12596 (+ 1 SD):

   Est.    S.E.    t val.       p
------- ------ -------- ------
  31.66   3.57     8.88   0.00
```

Statistical Model