An Introduction to Moderation

Overview

- Introduction to moderation
- Learn how to estimate and test moderation effects
- Learn how to estimate and test nonlinear effects

Job Demands-Resources (JDR) Theory: Our Example

The JDR Theory explains how job resources and demands influence one's engagement and burnout:

- Performance feedback moderates the relationship between cognitive demands and burnout
- Recovery opportunities moderate the relationship between emotional demands and burnout
- Work complexity moderates the relationship between performance feedback and engagement

Multiple Regression Model

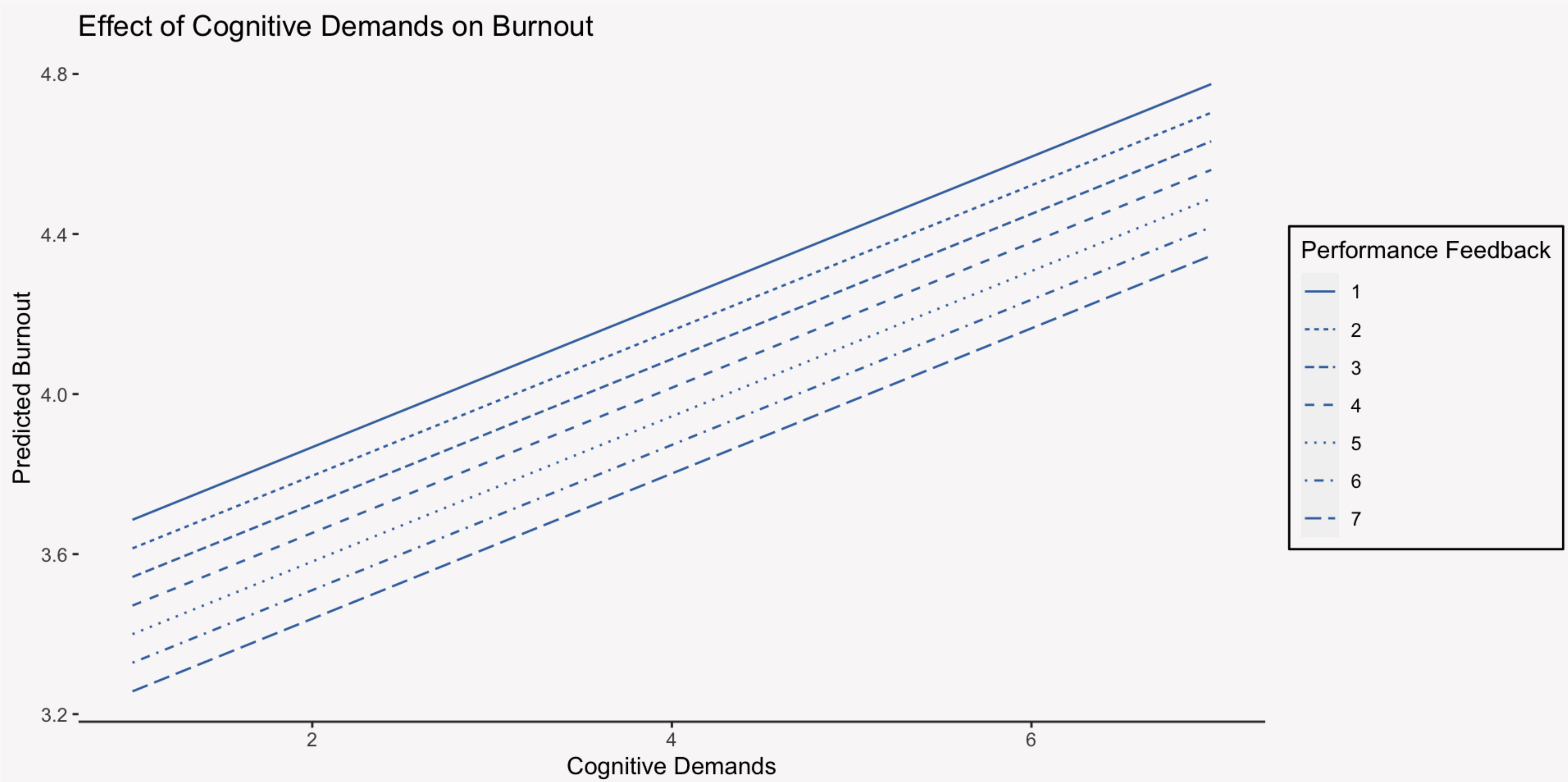
```r
mod_cog_demands <- lm(burnout ~ cog_demands + perf_feedback, data = data_1)
```

Interpreting a Linear Regression Model

```r
summary(mod_cog_demands)
```

```

Call:
lm(formula = burnout ~ cog_demands + perf_feedback, data = data_1)

Residuals:
    Min      1Q  Median      3Q     Max 
-3.1974 -0.8014  0.0171  0.6926  3.7431 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept)   3.57583    0.15481  23.098  < 2e-16 ***
cog_demands   0.18149    0.03113   5.830 7.49e-09 ***
perf_feedback -0.07148    0.02125  -3.364 0.000797 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9979 on 997 degrees of freedom
Multiple R-squared:  0.04263,    Adjusted R-squared:  0.04071 
```

The Unconditional Effect of Cognitive Demands on Burnout



What is Moderation (Statistical Interaction)?

Moderation or interaction, causal or otherwise, involves at least three variables in which the relationship between two variables, a focal predictor and outcome variable, changes depending on the value of a third variable, the moderating variable or moderator.

Some Moderation Jargon

When reading about moderation, you will likely come across several of these terms:

- Focal Predictor: Predictor who's relationship with the outcome is changing because of the moderator.
- Moderating Variable: Predictor that is altering (moderating) the relationship between the focal predictor and outcome.
- Conditional Effects: The effect of the focal predictor at a specific value of the moderator.
  - Main Effects
  - Lower & Higher Order Effects
  - Simple Slopes

Modeling an Interaction as a Cross-Product

To estimate an interaction effect, we create a new variable that is the product of the focal predictor and the moderator:

$$Y = \beta_0 + \beta_1 X + \beta_2 Z + \beta_3 XZ$$

$$Y = (\beta_0 + \beta_2 Z) + (\beta_1 + \beta_3 Z)X$$

Interaction Between Two Quantitative Variables

Below we fit a model looking at how two quantitative variables interact. Specifically, we are interested in determining if performance feedback moderates the relationship between cognitive demands and burnout:

```
mod_cog_demands_int <- lm(burnout ~ cog_demands + perf_feedback + cog_demands * perf_f

mod_cog_demands_int <- lm(burnout ~ cog_demands * perf_feedback, data = data_1)
```

Interpreting the Coefficients

- **Intercept**: Is the estimate of the outcome variable when the focal predictor and moderator both equal 0.
- **Slope of the Focal Predictor**: The estimated difference in the outcome variable between two cases that differ by 1 unit on the focal predictor and whose response to the moderator variable equals 0. It is the conditional effect of the focal predictor when the moderator equals 0.
- **Slope of the Moderator**: The estimated difference in the outcome variable between two cases that differ by 1 unit on the moderator and whose response to the focal predictor equals 0. It is the conditional effect of the moderator when the focal predictor equals 0.
- **Slope of the Interaction**: The estimated change in the conditional effect of the focal predictor as the moderator changes by one unit; The estimated change in the conditional effect of the moderator as the focal predictor changes by one unit.

Interpreting the Coefficients

```r
mod_cog_demands_int
```

```
Call:
lm(formula = burnout ~ cog_demands * perf_feedback, data = data_1)

Coefficients:
              (Intercept)                cog_demands
                  -0.3391                     1.1538
            perf_feedback  cog_demands:perf_feedback
                   0.8426                    -0.2265
```

Calculating the Conditional Effect of Cognitive Demands

We can calculate the conditional effect for Cognitive Demands **when the response to the Performance Feedback question is 7**:

$$Y = (-.34 \; + \; .84 \times 7) \; + (1.15 \; + \; -.23 \times 7)X$$

$$Y = 5.54 \; + \; -.46X$$

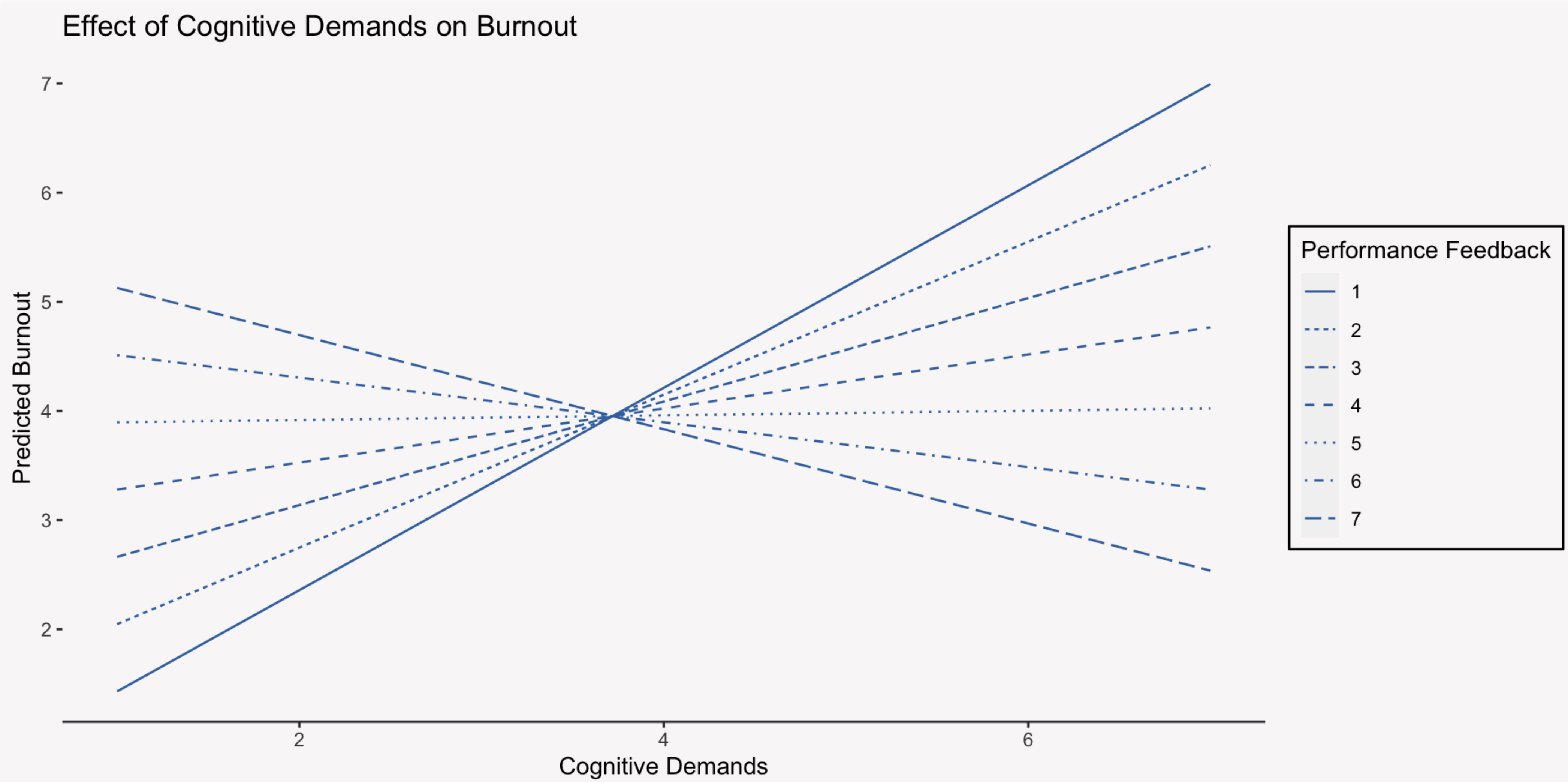
Calculating the Conditional Effect of Cognitive Demands

We can calculate the conditional effect for Cognitive Demands **when the response to the Performance Feedback question is 1**:

$$Y = (-.34\ +\ .84 \times 1)\ + (1.15\ +\ -.23 \times 1)X$$

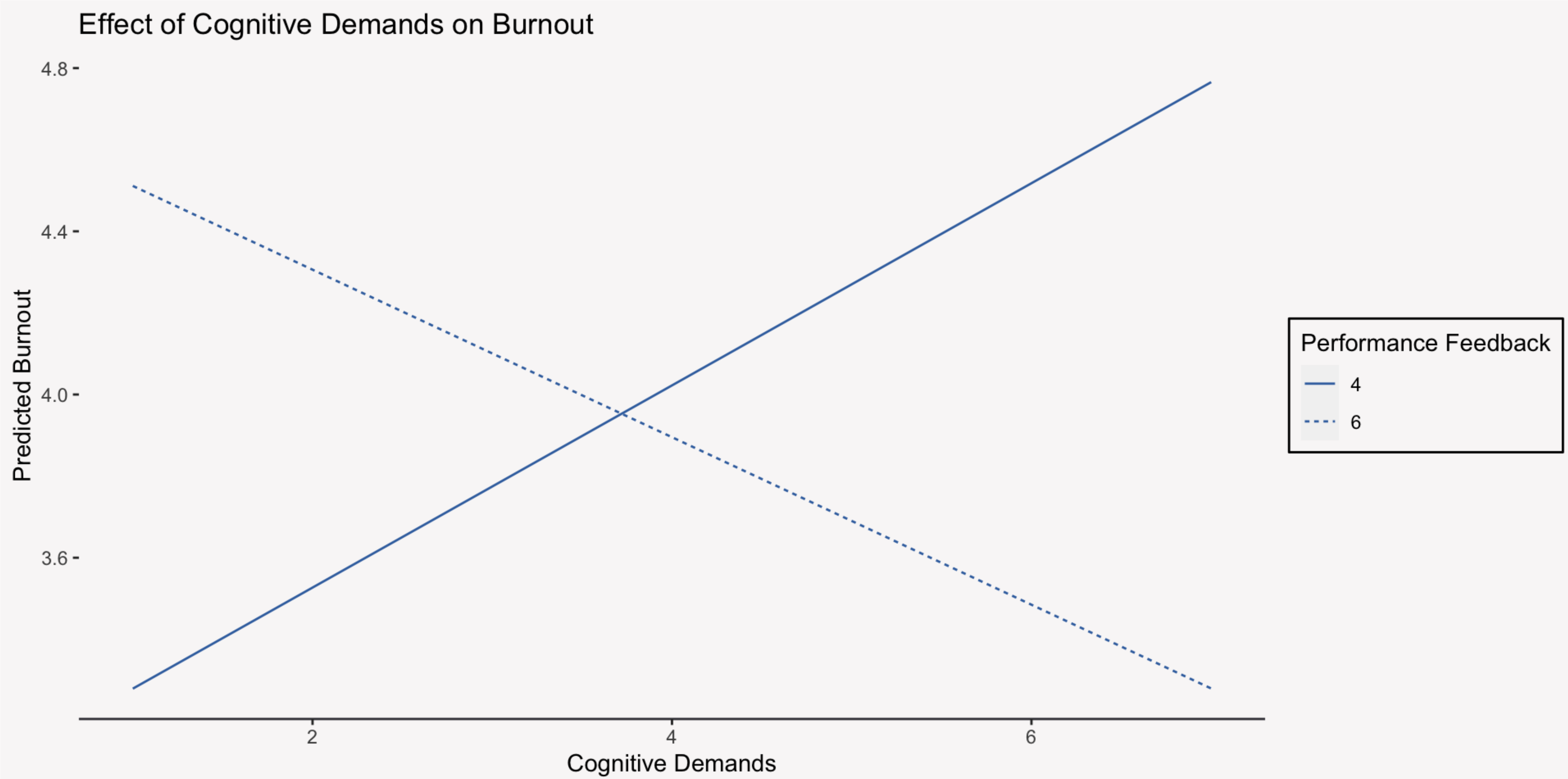
$$Y = .50\ +\ .92X$$

Plotting the Conditional Effect of Cognitive Demands on Burnout



Tips on Plotting the Interaction Effect

If you do not want to plot the conditional effect of the focal predictor at every level of the moderator, then you can choose the 36th, 50th, and 84th percentile of the moderator:



Testing the Interaction Effect

To determine if two variables interact with one another, we can use the p-value of the interaction regression slope to determine if the effect is significantly different than 0, or we can construct a 95% confidence interval around the regression slope and see if 0 is outside of the interval.

Testing the Interaction Effect: Statistical Test

```r
summary(mod_cog_demands_int)
```

```

Call:
lm(formula = burnout ~ cog_demands * perf_feedback, data = data_1)

Residuals:
    Min      1Q  Median      3Q     Max 
-3.2701 -0.6116  0.0199  0.5999  3.1046 

Coefficients:
                          Estimate Std. Error t value Pr(>|t|)    
(Intercept)               -0.33909    0.35570  -0.953    0.341    
cog_demands                1.15384    0.08579  13.450   <2e-16 ***
perf_feedback              0.84261    0.07842  10.744   <2e-16 ***
cog_demands:perf_feedback -0.22652    0.01880 -12.048   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9327 on 996 degrees of freedom
```

Testing the Interaction Effect: Confidence Interval

```r
confint(mod_cog_demands_int)
```

```
                                 2.5 %     97.5 %
(Intercept)                 -1.0370943  0.3589094
cog_demands                  0.9854904  1.3221892
perf_feedback                0.6887085  0.9965023
cog_demands:perf_feedback -0.2634147 -0.1896266
```

Interaction Between a Quantitative and Binary Variable

Below we estimate a model where a binary variable, recovery opportunities, moderates a quantitative variable, emotional demands:

```r
mod_emot_demands_int <- lm(burnout ~ emot_demands * rec_opp_factor, data = data_2)
```

Interpreting the Regression Coefficients

- **Intercept**: The estimated value of the outcome when both the focal predictor and moderator equal 0.
- **Slope of the Focal Predictor (Quantitative)**: The conditional effect of the focal predictor for the reference group (when the moderator equals 0).
- **Slope of the Moderator (Binary Variable)**: The conditional effect of the moderator when the focal predictor equals 0.
- **Slope of the Interaction**: The estimated change in the conditional effect of the focal predictor as the moderator changes by one unit.

Interpreting the Regression Coefficients

```r
mod_emot_demands_int
```

```

Call:
lm(formula = burnout ~ emot_demands * rec_opp_factor, data = data_2)

Coefficients:
                  (Intercept)                   emot_demands  
                       3.8469                         0.2194  
            rec_opp_factorYes  emot_demands:rec_opp_factorYes  
                       0.1653                        -0.2858  
```

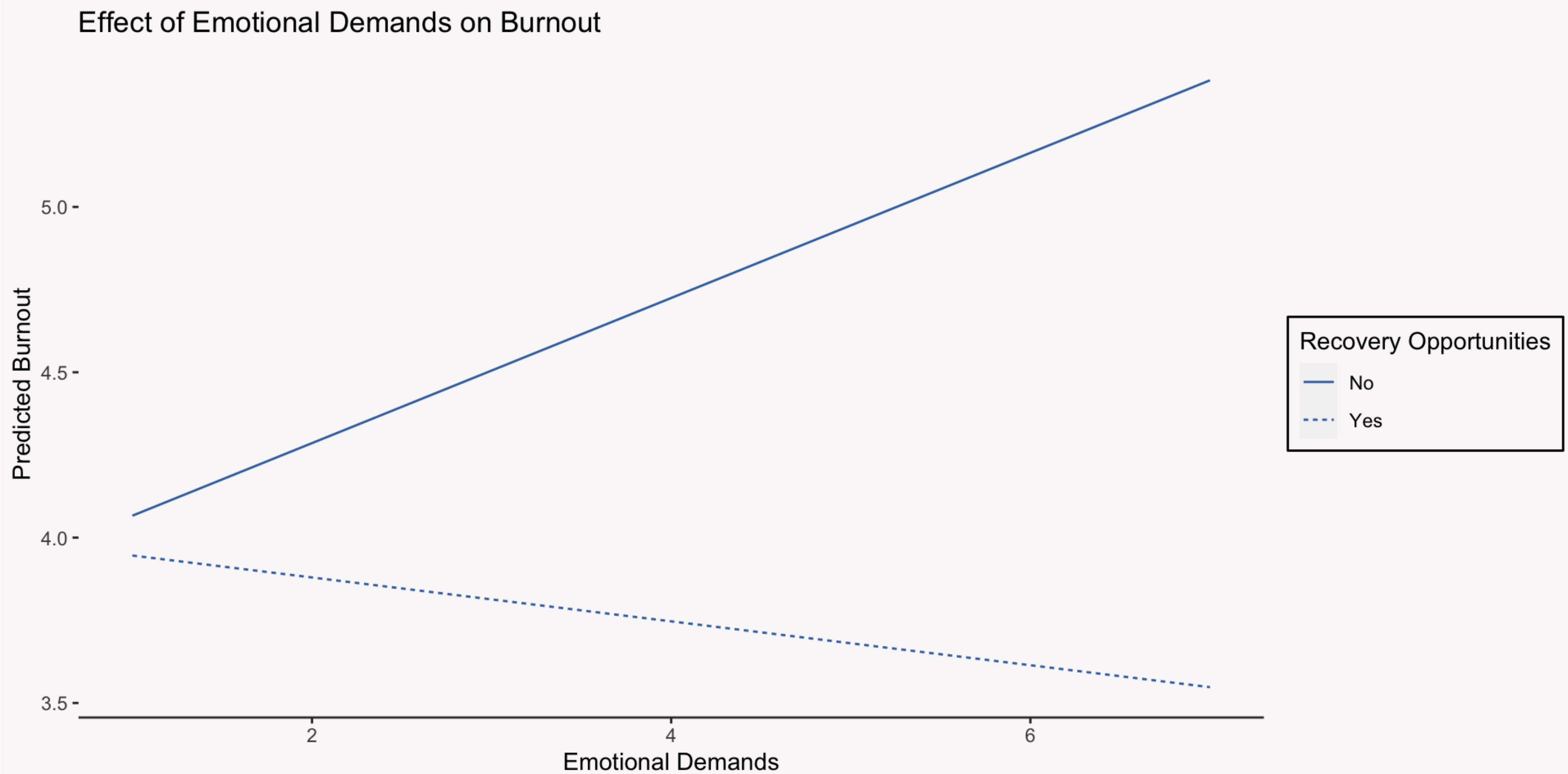
Calculating the Conditional Effects of Emotional Demands

We can calculate the conditional effect of emotional demands for the group that receives recovery opportunities as:

$$Y = (3.85 + .17 \times 1) \ + \ (.22 \ + \ - .29 \times 1)X$$

$$Y = 4.02 \ + \ - .07X$$

Plotting the Conditional Effects of Emotional Demands



Testing the Interaction Effect: Statistical Test

```r
summary(mod_emot_demands_int)
```

```

Call:
lm(formula = burnout ~ emot_demands * rec_opp_factor, data = data_2)

Residuals:
    Min      1Q  Median      3Q     Max 
-2.7468 -0.7246  0.1868  0.3195  3.2532 

Coefficients:
                              Estimate Std. Error t value Pr(>|t|)    
(Intercept)                    3.84691    0.23799  16.164  < 2e-16 ***
emot_demands                   0.21943    0.05692   3.855 0.000123 ***
rec_opp_factorYes              0.16529    0.28035   0.590 0.555606    
emot_demands:rec_opp_factorYes -0.28577    0.06704  -4.263 2.21e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9448 on 996 degrees of freedom
```

Interaction Between a Quantitative and Multicategorical Variable

Below we estimate a model where a multicategorical variable, work complexity, moderates a quantitative variable, performance feedback:

```
mod_perf_feed_int <- lm(engage ~ perf_feedback*work_complex)
```

Regression Model for Moderation with a Multicategorical Variable

$$Y = \beta_0 + \beta_1 X + \beta_2 Z_{\text{Low}} + \beta_3 Z_{\text{Mod.}} + \beta_4 X Z_{\text{Low}} + \beta_5 X$$

Interpreting the Regression Coefficients

The coefficients have similar interpretations as the previous models, but we just have a few more because we need multiple coefficients to capture the multicategorical variable:

```r
summary(mod_perf_feed_int)$coefficients[, 1:2] |> round(3)
```

```
                                Estimate Std. Error
(Intercept)                       -1.119      0.287
perf_feedback                      0.613      0.064
work_complexlow                    0.946      0.549
work_complexmoderate               0.936      0.484
perf_feedback:work_complexlow     -0.746      0.120
perf_feedback:work_complexmoderate -0.505      0.110
```

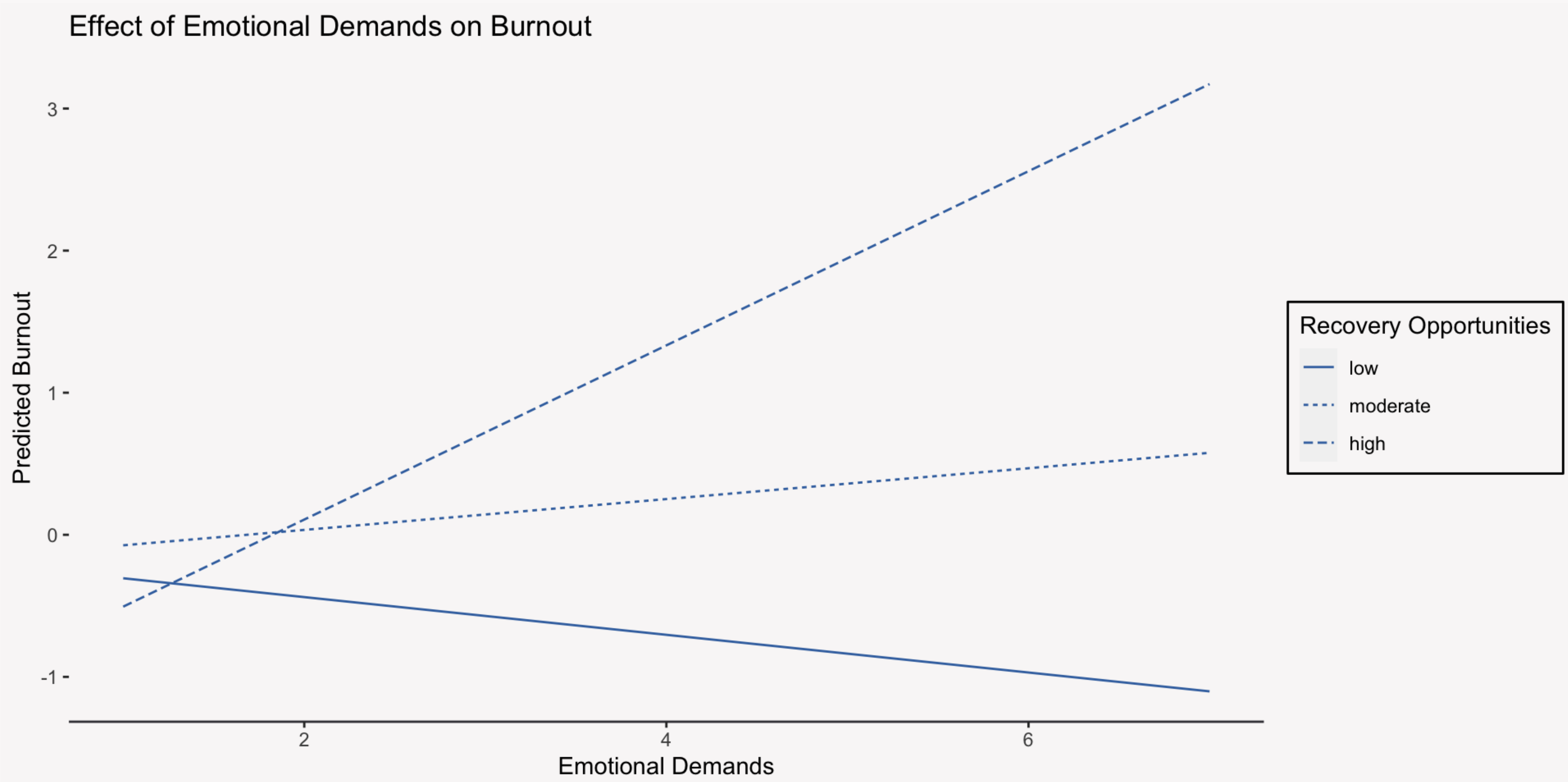
Calculating the Conditional Effects of Performance Feedback

We can calculate the conditional effect of performance feedback for the low work complexity group as:

$$Y = (-1.12 + .95 \times 1 + .94 \times 0) + (.61 + -.75 \times 1 + -.50 \times 0)X$$

$$Y = -.17 + -.14X$$

Plotting the Conditional Effects of Performance Feedback



Testing the Interaction Effect with a Multicategorical Moderator

Because a multicategorical moderator will have multiple interaction terms, referred to as a set, we cannot determine the significance of the overall set by looking at the significance of each individual interaction.

Instead, we will use something called an F-test to determine if including the set of interaction terms significantly increased the model's R^2 compared to the model without the set of interaction terms.

Testing the Interaction Effect with a Multicategorical Moderator

```r
# Fit a model without interactions and one with interactions
mod_no_int <- lm(engage ~ perf_feedback + work_complex, data = data_3)
mod_int <- lm(engage ~ perf_feedback * work_complex, data = data_3)

# Test the two models
anova(mod_no_int, mod_int)
```

```
Analysis of Variance Table

Model 1: engage ~ perf_feedback + work_complex
Model 2: engage ~ perf_feedback * work_complex
  Res.Df    RSS Df Sum of Sq      F   Pr(>F)
1    996 910.32
2    994 872.88  2    37.432 21.313 8.65e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The F-Test

The test we just conducted is referred to as an F-Test. It allows us to compare two models where one model is a subset of the other model, which is referred to as the full model.

It tests whether the R^2 of the full model is significantly larger than the R^2 of the reduced model:

$$R^2_{Full} - R^2_{Reduced} > 0$$

Testing Interactions vs Probing Interactions

So far we have only tested our interactions.

Testing an interaction is when you use either a statistical test or confidence interval to determine if the interaction's regression slope is statistically different from 0.

Probing an interaction is when you construct significance tests to determine at which values of the moderator the conditional effect (simple slope) of the focal predictor is significant.

Probing Interactions: Cognitive Demands X Performance Feedback

We will use the **pick a point** approach to probe our interactions. This is where you pick a specific value (point) of the moderator and test whether the conditional effect of the focal predictor is significant:

```
mod_min <- lm(burnout ~ cog_demands * I(perf_feedback - 1), data = data_1)
mod_median <- lm(burnout ~ cog_demands * I(perf_feedback - 4), data = data_1)
mod_max <- lm(burnout ~ cog_demands * I(perf_feedback - 7), data = data_1)
```

Probing Interactions: Cognitive Demands x Performance Feedback

```r
summary(mod_min)$coefficients |> round(3)
```

```
                                 Estimate Std. Error t value Pr(>|t|)
(Intercept)                         0.504      0.283   1.778    0.076
cog_demands                         0.927      0.068  13.557    0.000
I(perf_feedback - 1)                0.843      0.078  10.744    0.000
cog_demands:I(perf_feedback - 1)   -0.227      0.019 -12.048    0.000
```

Probing Interactions: Cognitive Demands x Performance Feedback

Conditional Effect	Estimate	Std. Error	t value	p-value
Min	0.927	0.068	13.557	0
Median	0.248	0.030	8.366	0
Max	-0.432	0.059	-7.364	0

Guidance for Modeling Interactions

- **Always** keep the lower order terms in the model even if they are not significant.
- You **do not** have to mean center your focal predictor and moderator before creating the interaction term.
- **Do not** artificially categorize a quantitative variable to make an interaction term.

Thinking of Nonlinear Effects as Interactions

Nonlinear effects occur whenever the relationship between a focal predictor and an outcome **changes at different values of the focal predictor**.

In a way, this is similar to the idea of moderation where X would be both the focal predictor and moderator.

Using Polynomial Regression to Model Nonlinear Effects

We can use linear regression to approximate most nonlinear functions by creating higher order (e.g. squaring, cubing) terms from the focal predictor:

$$Y = \beta_0 + \beta_1 X + \beta_2 X^2 + \beta_3 X^3$$

An Example of Polynomial Regression

We will estimate the curvilinear (nonlinear) relationship between **procrastination** and **creativity**:

```r
mod <- lm(creativity ~ procrastination + I(procrastination^2))

summary(mod)$coefficients
```

```
                       Estimate  Std. Error   t value      Pr(>|t|)
(Intercept)          7.32168766 0.159464662  45.91417 3.431270e-248
procrastination      0.52784158 0.034775741  15.17844  5.707781e-47
I(procrastination^2) -0.02598866 0.001601668 -16.22599  1.016472e-52
```

Plotting the Nonlinear Relationship