Online Supplement for Manuscript titled "Assessing and Interpreting Interaction Effects: A Reply to Vancouver, Carlson, Dhanani, & Colton (2021)"

Problems with Vancouver et al.'s Simulation

Vancouver et al. argued that ΔR^2 is problematic for estimating the size of interaction effects based on a Monte Carlo simulation that sought to estimate "true scores" for Y based on a "conceptual" model for the pure multiplying/interacting (i.e., joint) effects of X and Z on Y:

$$\text{Y} = \text{XZ} \qquad \text{(S1)}$$

They then analyzed the data using the following equations:

$$\text{Y} = \text{b}_0 + \text{b}_1\text{X} + \text{b}_2\text{Z} + \text{e} \qquad \text{(S2)}$$
$$\text{Y} = \text{b}_0 + \text{b}_1\text{X} + \text{b}_2\text{Z} + \text{b}_3\text{XZ} + \text{e} \qquad \text{(S3)}$$

Using S2 and S3 they computed $\Delta R^2 = .01$ and concluded that it provided a severe underestimate of the interaction effect size. However, their simulation equated data that were generated using a cross-product model (i.e., Y=XZ) with data generated using a pure multiplying model (i.e., Y=X×Z). The cross-product is **not** equivalent to the interaction effect (Aiken & West, 1993; Arnold, 1991; Arnold & Evans, 1979; Cohen et al., 2003). As LeBreton et al. (2013) summarized, "...the cross-product term ([XZ]) is a tainted, messy predictor containing information about three different statistical effects: (a) the main effect of [X], (b) the main effect of [Z], and (c) the [X×Z] interaction effect" (p. 454). To accurately simulate data that are only a function of a pure multiplying model, one must isolate the interaction effect and use it to simulate the data.

Correcting and Expanding Vancouver et al.'s Simulation

We corrected and also extended Vancouver et al.'s simulation in several ways. First, we simulated Y in a manner identical to Vancouver et al., which we denoted Y.XZ to clarify that what they referred to as a pure multiplying model was actually based on a *cross-product model.* Next, we simulated three additional versions of Y based on: (a) pure adding model (Y.ADD), (b) pure multiplying model (Y.XZ_{res}), and (c) combined adding-multiplying model (Y.COMB). To simulate a pure multiplying model, Y.XZ_{res}, it was necessary to sanitize the XZ cross-product of the additive effects of X and Z. This was accomplished by regressing XZ onto X+Z. The residuals from that analysis are equivalent to the "pure multiplying" effect (i.e., X×Z; LeBreton et al., 2013):

$$\text{XZ} = \text{b}_0 + \text{b}_1(\text{X}) + \text{b}_2(\text{Z}) + XZ_{res} \qquad \text{(S4)},$$

where XZ_{res} is equal to the interaction effect (X×Z). Finally, we also verified results using observed data that were estimated by introducing measurement error. The next section of the supplement summarizes the steps (and R code) we used in our simulation.

STEP 1: Compute True Scores for X, Z, and XZ

```
set.seed(1234) # Allows user to perfectly replicate simulation
n=100000; pop.mn=c(0.500,0.500)
pop.sd=c(0.100,0.100); pop.var = pop.sd^2
data <- data.frame(MASS::mvrnorm(n=n,mu=pop.mn,Sigma=diag(pop.var),empirical=TRUE))
names(data)=c("X", "Z") # Assigned names to the two variables
data$XZ <- data$X*data$Z # Compute the cross-product term
```

STEP 2: Compute the Interaction Effect By Residualizing the Cross-Product

```
INTER <- lm(XZ~X+Z, data= data)
data$INTER <- INTER$residuals
```

STEP 3: Generate Population Values for Y

To generate $Y.XZ_{res}$ with distributional characteristics consistent with Vancouver et al.'s Y.XZ, we rescaled $Y.XZ_{res}$ by converting it to a z-score, multiplying the z-score by .10 and adding .50.

```
data$Y.XZ  <- data$XZ
data$Y.ADD  <- data$X + data$Z
data$Y.XZres <- data$INTER
data$Y.XZres <- scale(data$INTER)*0.10+0.50
data$Y.COMB <-data$X + data$Z + data$XZ
```

STEP 4: Compute Observed Scores (x, z, y, xz) and Examine Correlations

Per Vancouver et al., observed scores (x, z, y, xz) were simulated by adding measurement error to the population true scores. We adjusted the error added to $Y.XZ_{res}$, Y.ADD, and Y.COMB to yield similar levels of criterion reliability to Vancouver et al.

```
data$x       <- data$X + rnorm(1000, 0.00, 0.05)
data$z       <- data$Z + rnorm(1000, 0.00, 0.05)
data$xz      <- data$x*data$z
data$y.xz    <- data$Y.XZ + rnorm(1000, 0.00, 0.05)
data$y.add   <- data$Y.ADD + rnorm(1000, 0.00, 0.10)
data$y.xzres <- data$Y.XZres + rnorm(1000, 0.00, 0.07)
data$y.comb <- data$Y.COMB + rnorm(1000, 0.00, 0.15)
```

```
##         X    Z    XZ   INTER Y.XZ Y.ADD Y.XZr Y.COM x    z    xz
## X       1.00
## Z       0.00 1.00
## XZ      0.70 0.70 1.00
## INTER   0.00 0.00 0.14 1.00
## Y.XZ    0.70 0.70 1.00 0.14  1.00
## Y.ADD   0.71 0.71 0.99 0.00  0.99 1.00
## Y.XZres 0.00 0.00 0.14 1.00  0.14 0.00  1.00
## Y.COMB  0.71 0.71 1.00 0.05  1.00 1.00  0.05  1.00
## x       0.90 0.00 0.63 0.00  0.63 0.64  0.00  0.64  1.00
## z       0.00 0.89 0.63 0.00  0.63 0.63  0.00  0.63  0.00 1.00
## xz      0.63 0.63 0.90 0.13  0.90 0.89  0.13  0.89  0.69 0.70 1.00
## y.xz    0.58 0.59 0.84 0.12  0.84 0.83  0.12  0.83  0.54 0.53 0.76
## y.add   0.57 0.57 0.80 0.00  0.80 0.81  0.00  0.81  0.52 0.51 0.73
## y.xzres 0.00 0.00 0.11 0.82  0.11 0.00  0.82  0.04  0.00 0.01 0.11
## y.comb  0.58 0.58 0.82 0.04  0.82 0.82  0.04  0.82  0.52 0.53 0.74
##         y.xz y.add y.xzr y.cmb
## y.xz    1.00
## y.add   0.65 1.00
## y.xzres 0.09 0.01  1.00
## y.comb  0.68 0.66  0.03  1.00
```

True-Score Correlations
Our simulation reproduced the relationships reported in Vancouver et al:

- X and Z were orthogonal.
- Y.XZ had strong bivariate (non-joint) correlations with X and Z.
- XZ had strong bivariate correlations with its components, X and Z.
- Y.XZ had a perfect correlation with the cross-product, XZ.

Our simulation clarified the problems with the Vancouver et al. simulation:

- Y.XZ was correlated .99 with Y.ADD (i.e., the pure adding model), suggesting Vancouver et al.'s cross-product model was driven almost entirely by the additive (i.e., non-joint) effects of X and Z.

Our simulation correctly generated scores for Y based on the pure multiplying model:

- $Y.XZ_{res}$ was correlated .14 with the cross-product and was uncorrelated with X and Z. This suggests that $Y.XZ_{res}$ is driven **solely** by the multiplying effect of X and Z.

Observed-Score Correlations
The observed score correlations were consistent with the construct score correlations, but were attenuated due to measurement error. The correlations between observed scores and true scores were equal to the reliability index. If we square those correlations, we obtain estimates of the reliability coefficients for our variables: $r_{xx} = 0.81$, $r_{zz} = 0.8$, $r_{yy.add} = 0.66$, $r_{yy.xzres} = 0.67$, $r_{yy.comb} = 0.67$, $r_{yy.xz} = 0.7$.

STEP 5: Conduct Test of Moderated Multiple Regression

Moderated multiple regression (MMR) is a two-step test. The interaction effect is estimated as the difference in R^2 values obtained after estimating S2 and S3 (i.e., ΔR^2). The null hypothesis significance test for ΔR^2 is identical to that for b_3. If the null hypothesis is rejected, then S2 is rejected and S3 is retained as the correct model. The functional form of the interaction is determined by graphing simple slopes. If the interaction effect is non-significant, then S3 is rejected and S2 is retained.[1]

We denoted results based on population data using uppercase letters and sample data using lowercase letters. Labels for each set of results first identified the true data generating mechanism used to compute Y/y (e.g., Y.XZ, Y.ADD, Y.XZres, Y.COMB, y.xz, y.add, y.xzres, y.comb) before indicating the regression model used to predict Y/y as either additive (.ADD/.add) or combined (.COMB/.comb). For example, Y.ADD.ADD denotes: a) population values are being tested, b) the data-generating function for Y was pure adding (i.e., Y = X + Z) and, c) the results are based on the additive predictor model (i.e., X+Z).

```
# Testing Models Using True Scores (Constructs)
Y.ADD.ADD   <- lm(Y.ADD~X+Z, data=data) # Step 1 of MMR
Y.ADD.COMB  <- lm(Y.ADD~X+Z+XZ, data=data) # Step 2 of MMR
Y.XZres.ADD  <- lm(Y.XZres~X+Z, data=data) # Step 1 of MMR
Y.XZres.COMB <- lm(Y.XZres~X+Z+XZ, data=data) # Step 2 of MMR
Y.COMB.ADD  <- lm(Y.COMB~X+Z, data=data) # Step 1 of MMR
Y.COMB.COMB <- lm(Y.COMB~X+Z+XZ, data=data) # Step 2 of MMR
Y.XZ.ADD    <- lm(Y.XZ~X+Z, data=data) # Step 1 of MMR
Y.XZ.COMB   <- lm(Y.XZ~X+Z+XZ, data=data) # Step 2 of MMR

# Testing Models Using Observed Scores
y.add.add  <- lm(y.add~x+z, data=data)
y.add.comb <- lm(y.add~x+z+xz, data=data)
y.xzres.add  <- lm(y.xzres~x+z, data=data)
y.xzres.comb <- lm(y.xzres~x+z+xz, data=data)
y.comb.add  <- lm(y.comb~x+z, data=data)
y.comb.comb <- lm(y.comb~x+z+xz, data=data)
y.xz.add  <- lm(y.xz~x+z, data=data)
y.xz.comb <- lm(y.xz~x+z+xz, data=data)
```

STEP 6: Prepare the Data for Graphing

In this step, we prepared the data necessary to graph the models from Step 5. This will help to visually depict whether MMR yielded results consistent with the data-generating functions for Y. First, we estimated means and SDs for X, Z, x, and z. We used those values to estimate high and low values of X, Z, x, and z. Then we added scores for the cross-product terms. Finally, we estimated predicted values for the different versions of Y/y.

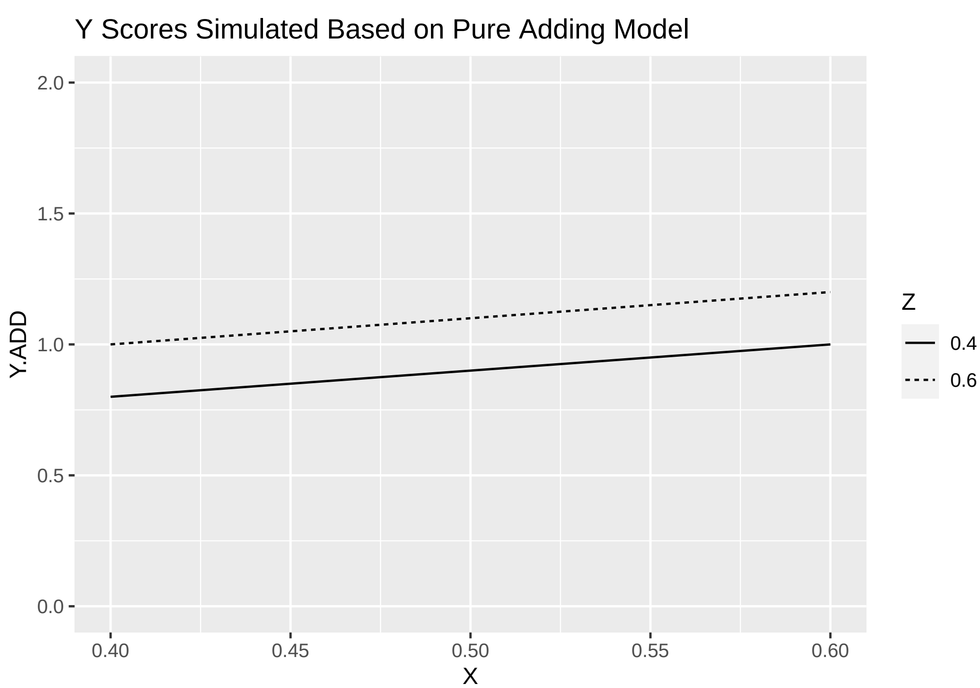
[1] Code for generating output based on observed scores is included in rmarkdown file, but results not included in pdf.

STEP 7: Results of Simulation

Data Generating Function for Y = Pure Adding Model

When the data generating function for Y was based on a pure adding model, MMR correctly rejects the interaction and correctly partitions R^2 across lower-order ($R^2 = 1.00$) and interaction ($R^2 = 0.00$) effects. Graphing the results of this model reveals a strong additive effect for X, which is consistent with the data generating function.

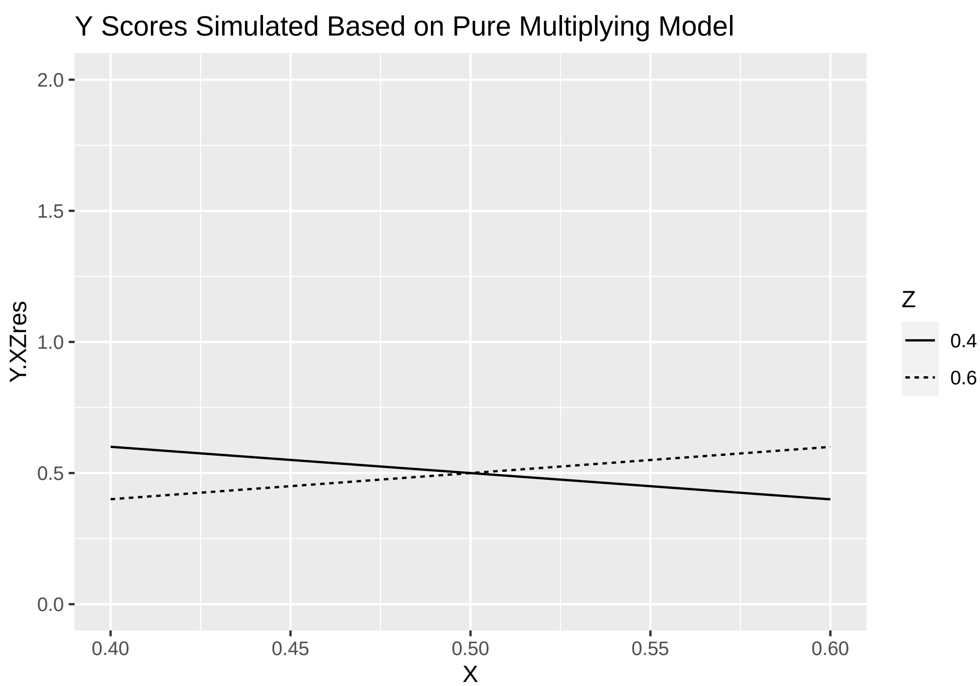
```
##
## =======================================
##                       Y.ADD
##           ----------------------------
##              Step 1         Step 2
##               (1)            (2)
## ---------------------------------------
## Constant   -0.000***      -0.000***
##             (0.000)        (0.000)
## X           1.000***       1.000***
##             (0.000)        (0.000)
## Z           1.000***       1.000***
##             (0.000)        (0.000)
## XZ                         -0.000
##                            (0.000)
## ---------------------------------------
## R2           1.000          1.000
## =======================================
## Note:     *p<0.1; **p<0.05; ***p<0.01
```



Data Generating Function for Y = Pure Multiplying Model

When the data generating function for Y was the pure multiplying model, MMR correctly detects the presence of an interaction and correctly partitions R^2 across lower-order ($R^2 = .00$) and interaction ($R^2 = 1.00$) effects. The assignment of the entire R^2 to the multiplicative effect is consistent with the disordinal interaction effect used to generate the data.

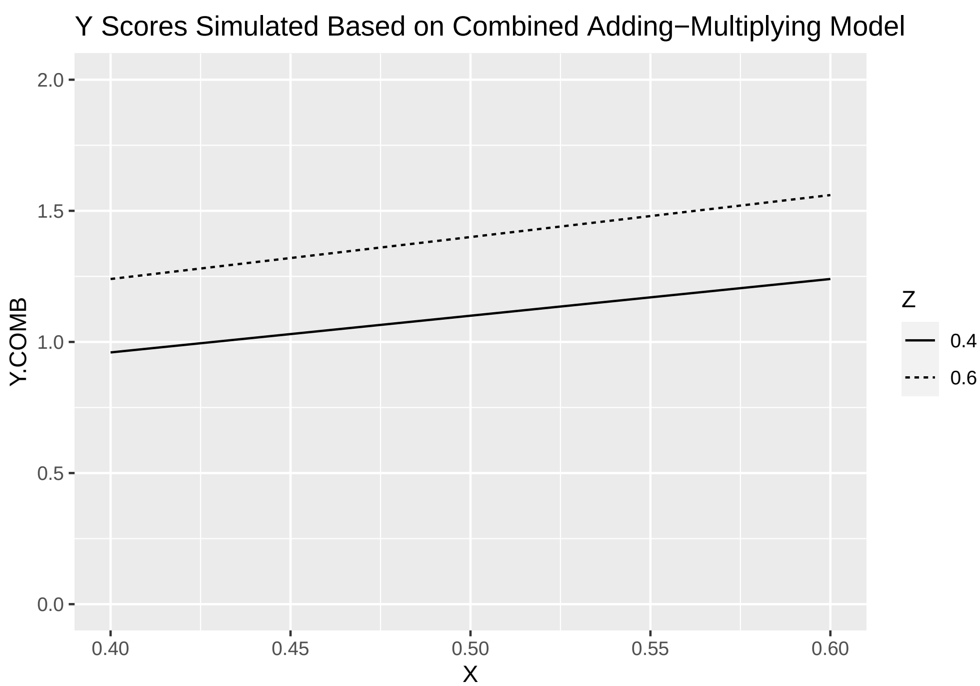
```
##
## ======================================
##                    Y.XZres
##             ----------------------------
##                 Step 1        Step 2
##                  (1)           (2)
## --------------------------------------
## Constant    0.500***      3.000***
##             (0.002)        (0.000)
## X            -0.000       -5.001***
##             (0.003)        (0.000)
## Z            -0.000       -4.999***
##             (0.003)        (0.000)
## XZ                        9.998***
##                            (0.000)
## --------------------------------------
## R2           0.000          1.000
## ======================================
## Note:     *p<0.1; **p<0.05; ***p<0.01
```



Data Generating Function for Y = Combined Adding-Multiplying Model

When the data generating function for Y was the combined adding-multiplying model, MMR correctly detects the interaction and partitions R^2 across additive ($R^2 = 0.998$) and interactive ($R^2 = 0.002$) effects. The ΔR^2 = .002 indicates trivial improvement in prediction. The small increment in R^2, coupled with the strong additive effects, is consistent with the ordinal interaction effect used to generate the data.

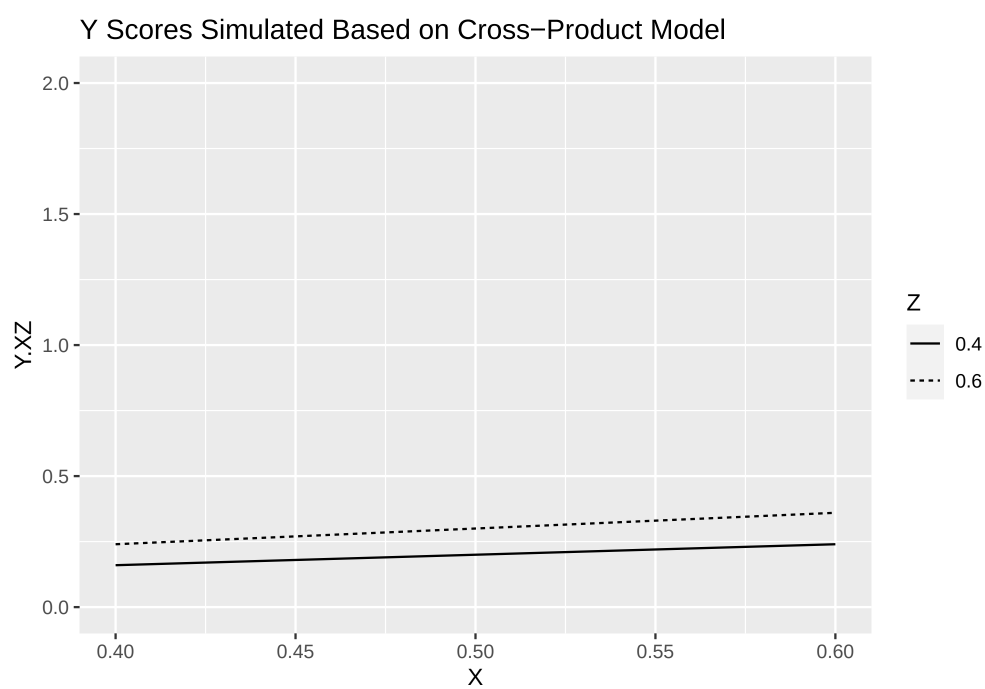
```
##
## =======================================
##                     Y.COMB
##             ---------------------------
##                 Step 1        Step 2
##                  (1)           (2)
## ---------------------------------------
## Constant   -0.250***      0.000***
##             (0.0002)       (0.000)
## X            1.500***      1.000***
##             (0.0003)       (0.000)
## Z            1.500***      1.000***
##             (0.0003)       (0.000)
## XZ                         1.000***
##                            (0.000)
## ---------------------------------------
## R2            0.998         1.000
## =======================================
## Note:     *p<0.1; **p<0.05; ***p<0.01
```



Data Generating Function for Y = Cross-Product Model

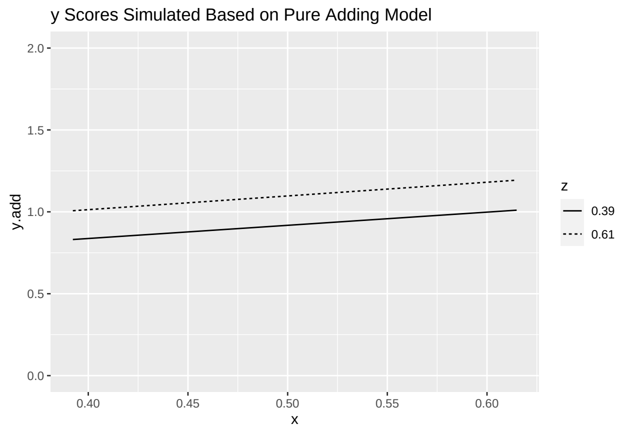
When the data generating function for Y was the cross-product model, MMR correctly detects the presence of an interaction and correctly partitions R^2 across additive ($R^2 = .98$) and interactive ($R^2 = .02$) effects. The $\Delta R^2 = .02$ indicates the interaction provides only a small improvement in prediction. The small increment in R^2, coupled with the strong additive effects, is consistent with the ordinal interaction effect used to generate the data. And, this interaction is consistent with the ordinal interaction effect simulated by Vancouver et al.

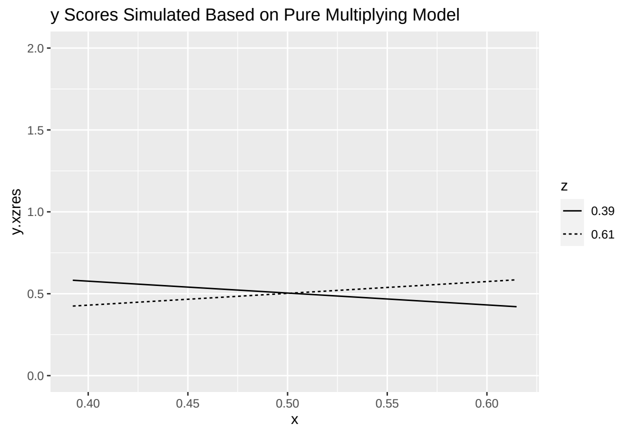
```
##
## ======================================
##                      Y.XZ
##          ----------------------------
##             Step 1        Step 2
##              (1)           (2)
## --------------------------------------
## Constant   -0.250***      -0.000***
##             (0.0002)       (0.000)
## X           0.500***      0.000***
##             (0.0003)       (0.000)
## Z           0.500***      0.000***
##             (0.0003)       (0.000)
## XZ                        1.000***
##                            (0.000)
## --------------------------------------
## R2           0.980          1.000
## ======================================
## Note:     *p<0.1; **p<0.05; ***p<0.01
```

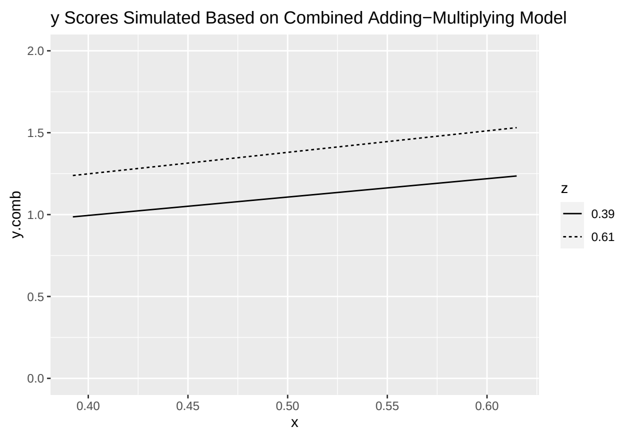


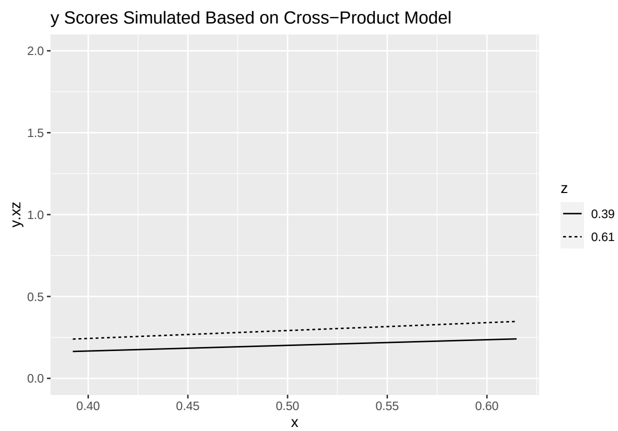
STEP 8: Verify Simulation Results Using Observed Data

In this step, we graphed the models using observed scores. All graphs were consistent with previous findings based on true scores. Thus, no substantive interpretations are included, only the graphs.









Summary

The results of our simulation serve to verify several important characteristics of MMR:

1) It correctly detects the presence or absence of an interaction effect,
2) It provides an accurate estimate of the magnitude of the interaction effect, and
3) It provides the information necessary to ascertain the functional form of the interaction effect.