Moderation

Fundamentals:

- Moderation refers to a change in the relationship between an independent variable and a dependent variable, depending on the level of a third variable, termed the moderator variable. Moderating effects are also referred to as interaction and conditioning effects.

* For two continuous variables, moderation means that the slope of the relationship between the independent and dependent variable varies (i.e., increases or decreases) according to the level of the moderator variable.

* For a continuous independent variable and a categorical moderator variable, moderation means that the slope of the relationship between the independent and dependent variable differs across the groups represented by the categorical moderator variable.

* For a categorical independent variable and a continuous moderator variable, moderation means that the differences between the group means represented by the levels of the categorical independent variable differ according to the level of the moderator variable.

* For two categorical variables, moderation means that the difference between the group means for the categorical independent variable differ depending on group membership on the moderator variable.

- When the predictor and moderator variables are continuous, a single product is needed to capture the moderating effect. When one variable is continuous and the other is categorical, the required number of product terms is \( g - 1 \), where \( g \) equals the number of groups represented by the categorical variable. When both variables are categorical, the required number of product terms is \( (g_1 - 1)(g_2 - 1) \), where \( g_1 \) and \( g_2 \) are the number of groups represented by the two categorical variables.

- Interactions can range up to the \( k \)th order, where \( k \) represents the number of variables on the right side of the equation. The \( k \)th order interaction is tested using the \( k \)th order product term, controlling for the first-order terms and all products formed by dropping one or more terms from the \( k \)th order product. Thus, a three-way interaction is testing using \( X_1X_2X_3 \) controlling for \( X_1, X_2, X_3, X_1X_2, X_1X_3, \) and \( X_2X_3 \). Analogously, a four-way interaction is testing using \( X_1X_2X_3X_4 \) controlling for \( X_1, X_2, X_3, X_4, X_1X_2, X_1X_3, X_1X_4, X_2X_3, X_2X_4, X_3X_4, X_1X_2X_3, X_1X_2X_4, X_1X_3X_4, \) and \( X_2X_3X_4 \).
- Moderation is represented not by a product term (or set of product terms) itself, but rather by a product term from which all lower-order terms constituting the product have been partialed.

- A curvilinear effect is a special case of moderation in which the relationship between an independent and dependent variable varies according to the level of the independent variable itself. For example, a U-shaped relationship between X and Y indicates a negative slope at low levels of X, a flat slope at moderate levels of X, and a positive relationship at high levels of X.

- Moderator variables should be distinguished from mediator variables, in that the latter transmit the effect of the independent variable on the dependent variable. For example, $X_2$ is a mediator of the effect of $X_1$ on $Y$ if $X_1$ influences $X_2$, which in turn influences $Y$. In contrast, $X_2$ is a moderator of the effect of $X_1$ on $Y$ if the relationship between $X_1$ and $Y$ varies according to the level of $X_2$. Although these processes are fundamentally different, the terms mediator and moderator are often confused in the literature.

**Methods of Analysis:**

- Moderation can be tested by forming product terms using the variables that constitute the interaction and estimating the increment in variance explained by the product terms after the lower-order terms they contain have been controlled. However, the form and interpretation of these terms can vary, depending on whether the constituent variables are continuous or categorical and on the number of levels of the categorical variable.

- **Two Categorical Variables:** For two dichotomous variables (i.e., categorical variables with only two levels), moderation is estimated as follows:

$$Y = b_0 + b_1X_1 + b_2X_2 + b_3X_1X_2 + e$$

If $X_1$ and $X_2$ are dummy coded (i.e., 0, 1), then the product $X_1X_2$ will equal 1 when both $X_1$ and $X_2$ are 1 and will equal zero otherwise. Thus, $X_1X_2$ can be viewed as a third dummy variable that is coded 1 when $X_1$ and $X_2$ are both 1. If $X_1$ and $X_2$ are arranged in a 2 x 2 table, the coefficient on $X_1X_2$ (i.e., $b_3$) can be interpreted as the change in the mean differences of $Y$ predicted by $X_1$ as one moves from $X_2 = 0$ to $X_2 = 1$. This interpretation is illustrated by the following table:
The entries in this table represent the mean of Y for different combinations of X₁ and X₂. The difference in the mean of Y for X₁ = 0 and X₁ = 1 is 5 – 1 = 4 for X₂ = 0 and 7 – 3 = 4 for X₂ = 1, both of which equal 4. Thus, there is no interaction between X₁ and X₂. Compare this to the following table:

<table>
<thead>
<tr>
<th>X₁</th>
<th>X₂</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>7</td>
</tr>
</tbody>
</table>

In this case, the mean difference in Y for X₁ = 0 and X₁ = 1 is 5 – 1 = 4 for X₂ = 0 but is 9 – 3 = 6 for X₂ = 1. This represents an interaction between X₁ and X₂, because the effect of X₁ varies, depending on the level of X₂.

In the above table, the mean of Y when X₁ and X₂ are both 1 is higher than in the table where X₁ and X₂ do not interact. It makes intuitive sense that this type of interaction could be captured by a dummy variable that contrasts the cell where X₁ and X₂ are both 1 against the other three cells, as represented by the X₁X₂ product. However, due to the partialing that occurs in regression analysis, the X₁X₂ product will capture an interaction between X₁ and X₂ of any form.
This principle is illustrated by estimating the means in the above table with $X_1$ and $X_2$ but not $X_1X_2$ as predictors. Assuming each cell contains the same number of observations, the effect for $X_1$ is the average of the mean differences in $Y$ for $X_1 = 0$ and $X_1 = 1$ across levels of $X_2$. The difference was 4 when $X_2 = 0$ and 6 when $X_2 = 1$, and thus the average is 5. Under these conditions, the estimated means using $X_1$ and $X_2$ as predictors are:

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.5</td>
<td>5.5</td>
</tr>
<tr>
<td>1</td>
<td>3.5</td>
<td>8.5</td>
</tr>
</tbody>
</table>

The effect of $X_1$ is one unit too large when $X_2 = 0$ and one unit too small when $X_2 = 1$. These discrepancies produce the following pattern of mean residuals:

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.5</td>
<td>-0.5</td>
</tr>
<tr>
<td>1</td>
<td>-0.5</td>
<td>0.5</td>
</tr>
</tbody>
</table>

Likewise, the effect for $X_2$ is the average of the mean differences in $Y$ for $X_2 = 0$ and $X_2 = 1$ across levels of $X_1$. The difference was 2 when $X_1 = 0$ and 4 when $X_1 = 1$, and therefore the average is 3. Thus, the effect of $X_2$ is one unit too large when $X_1 = 0$ and one unit too small when $X_1 = 1$. Again, these discrepancies are consistent with the pattern of mean residuals shown above.
Although the pattern of mean residuals shown above does not correspond to the contrast represented by the $X_1X_2$ product, it corresponds to the partialed $X_1X_2$ product. This can be seen by residualizing $X_1X_2$ on $X_1$ and $X_2$, which yields the following values for $X_1X_2$ for the four combinations of $X_1$ and $X_2$:

<table>
<thead>
<tr>
<th>$X_1$</th>
<th>$X_2$</th>
<th>$-0.25$</th>
<th>$0.25$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>-0.25</td>
<td>0.25</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0.25</td>
<td>0.25</td>
</tr>
</tbody>
</table>

The residual values for $X_1X_2$ are half of the residual values for $Y$. Hence, after controlling for $X_1$ and $X_2$, the $X_1X_2$ product is perfectly correlated with the mean residuals for $Y$ and therefore will explain the variance in the means of $Y$ not explained by $X_1$ and $X_2$. The pattern of mean residuals for $Y$ is the same for any interaction between $X_1$ and $X_2$, and therefore the partialed $X_1X_2$ product will always capture the interaction between $X_1$ and $X_2$.

For categorical variables with more than two levels, an interaction is represented by all pairwise products between the dichotomous variables used to represent the two categorical variables. For example, assume that both categorical variables represent three groups, and that two groups for the first variable are represented by the dummy variables $X_1$ and $X_2$, and two groups for the second variable are represented by the dummy variables $X_3$ and $X_4$. An equation containing their interaction would be as follows:

$$Y = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 + b_5X_1X_3 + b_6X_1X_4 + b_7X_2X_3 + b_8X_2X_4 + e$$

The four product terms $X_1X_3$, $X_1X_4$, $X_2X_3$, and $X_2X_4$ represent the interaction between the two original categorical variables. In general, these terms should be tested and interpreted as a set. If a significant increment in $R^2$ is yielded by these terms, an interaction is indicated, and the nine means associated with the 3x3 classification of the two categorical variables can be compared to determine the form of the interaction.
One Continuous and One Categorical Variable: The interaction between a continuous and dichotomous variable is estimated as follows:

\[ Y = b_0 + b_1 X_1 + b_2 X_2 + b_3 X_1 X_2 + e \]

If \( X_2 \) is dichotomous and framed as the moderator variable, then \( b_3 \) can be interpreted as the difference in the slope of \( Y \) on \( X_1 \) for the two groups represented by \( X_2 \). This can be seen by treating \( X_2 \) as a dummy variable and substituting either 0 or 1 for \( X_2 \) in the preceding equation. For example, if \( X_2 = 0 \), this equation becomes:

\[ Y = b_0 + b_1 X_1 + e \]

Thus, \( b_1 \) represents the slope of \( Y \) on \( X_1 \) for the omitted group (i.e., the group for which \( X_2 = 0 \)). If \( X_2 = 1 \), the equation becomes:

\[ Y = (b_0 + b_2) + (b_1 + b_3)X_1 + e \]

These expressions give simple slopes relating \( X_1 \) to \( Y \) when \( X_2 \) equals 0 and 1, respectively. Note that, if \( b_3 = 0 \), the slope of \( Y \) on \( X_1 \) is the same for both groups (i.e., for \( X_2 = 0 \) and \( X_2 = 1 \)), indicating no interaction between \( X_1 \) and \( X_2 \).

If a categorical moderator variable has more than two levels, its interaction with a continuous variable is represented by the products of the continuous variable with the \( g - 1 \) dichotomous variables used to represent the categorical variable. For example, if the continuous variable is represented by \( X_1 \) and the categorical variable has three levels and is represented by the dummy variables \( X_2 \) and \( X_3 \), the equation containing their interaction is as follows:

\[ Y = b_0 + b_1 X_1 + b_2 X_2 + b_3 X_3 + b_4 X_1 X_2 + b_5 X_1 X_3 + e \]

The equation for the omitted group is found by substituting 0 for \( X_2 \) and \( X_3 \):

\[ Y = b_0 + b_1 X_1 + e \]

The equation for the group represented by \( X_2 \) (i.e., when \( X_2 = 1 \)) is:

\[ Y = (b_0 + b_2) + (b_1 + b_4)X_1 + e \]

The equation for the group represented by \( X_3 \) (i.e., when \( X_3 = 1 \)) is:

\[ Y = (b_0 + b_3) + (b_1 + b_5)X_1 + e \]
Thus, the difference between the slope for the omitted group and the group represented by $X_2$ is reflected by $b_4$, and the difference between the slope for the omitted group and the group represented by $X_3$ is reflected by $b_5$. A joint (i.e., omnibus) test of whether the slopes for the three groups differ from one another is performed by simultaneously testing the constraints $b_4 = 0$ and $b_5 = 0$. The slopes for the groups represented by $X_2$ and $X_3$ can be compared by testing the difference between $b_4$ and $b_5$.

- Two Continuous Variables: When the independent and moderator variables are continuous, their interaction is estimated as follows:

$$Y = b_0 + b_1 X_1 + b_2 X_2 + b_3 X_1 X_2 + e$$

The change in slope of $Y$ on $X_1$, given a unit change in $X_2$, is represented by $b_3$. Note that the interpretation of $b_3$ is symmetric, in that it also represents the change in slope of $Y$ on $X_2$, given a unit change in $X_1$.

The preceding equation can be rewritten to explicitly show how the slope and intercept of the equation relating $X_1$ to $Y$ varies according to $X_2$:

$$Y = (b_0 + b_2 X_2) + (b_1 + b_3 X_2)X_1 + e$$

This equation represents the simple slope relating $X_1$ to $Y$ at a given level of $X_2$ as the compound coefficient $b_1 + b_3 X_2$. Note that the slope of $Y$ on $X_1$ changes by $b_3$ for a unit change in $X_2$. This equation also shows that $b_1$ represents the slope of $Y$ on $X_1$ when $X_2 = 0$.

The same expression for the slope of $Y$ on $X_1$ is obtained by taking the partial derivative of $Y$ with respect to $X_1$ in the preceding equation:

$$\frac{\delta Y}{\delta X_1} = b_1 + b_3 X_2$$

This again shows that the slope of $Y$ on $X_1$ is not a constant, but instead varies according to the level of $X_2$, and that the slope of $Y$ on $X_1$ at the point $X_2 = 0$ is represented by $b_1$.

Higher-Order Interactions:

- Higher-order interactions can be tested by extending the procedures described previously. For example, the interaction between three continuous variables $X_1$, $X_2$, and $X_3$ can be tested with the following equation:
\[ Y = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_1X_2 + b_5X_1X_3 + b_6X_2X_3 + b_7X_1X_2X_3 + e \]

In this equation, the three-way interaction is represented by the partialed triple product \( X_1X_2X_3 \). Thus, \( b_7 \) can be interpreted as the change in the interaction between any two variables (e.g., \( X_1 \) and \( X_2 \)) according to the level of the third variable (e.g., \( X_3 \)). Note that the equation for the three-way interaction contains the three first-order terms and all interactions terms of lower order (i.e., the three two-way interactions). This procedure can be extended to interactions of higher order.

**Testing Simple Slopes:**

- In moderated regression, simple slopes refer to the relationship between the independent variable and the dependent variable at selected levels of the moderator variable. To illustrate, reconsider a regression equation in which the relationship between \( X_1 \) and \( Y \) is moderated by the dummy variable \( X_2 \):

\[ Y = b_0 + b_1X_1 + b_2X_2 + b_3X_1X_2 + e \]

When \( X_2 = 0 \), this equation reduces to:

\[ Y = b_0 + b_1X_1 + e \]

When \( X_2 = 1 \), the equation becomes:

\[ Y = (b_0 + b_2) + (b_1 + b_3)X_1 + e \]

In these two expressions, the coefficients on \( X_1 \) are simple slopes relating \( X_1 \) to \( Y \) at \( X_2 = 0 \) and \( X_2 = 1 \). When \( X_2 = 0 \), the slope relating \( X_1 \) to \( Y \) is simply \( b_1 \), and this coefficient can be tested in the usual manner. When \( X_2 = 1 \), the slope relating \( X_1 \) to \( Y \) is given by the expression \( b_1 + b_3 \), which can be tested using procedures for testing linear combinations of regression coefficients. For instance, a t-test can be conducted for \( b_1 + b_3 \) by dividing this sum by its standard error, which can be obtained by taking the square root of the variance of \( b_1 + b_3 \). This variance can be derived using the squares of the standard errors of \( b_1 \) and \( b_3 \) to represent their variances and by computing the covariance between \( b_1 \) and \( b_3 \) as the product of their correlation with their standard errors. The resulting expression can be written as follows:

\[ V(b_1 + b_3) = V(b_1) + V(b_3) + 2C(b_1,b_3) \]
Dividing $b_1 + b_3$ by the square root of this expression yields a t-test for the simple slope $b_1 + b_3$.

When $X_2$ is continuous, the simple slope relating $X_1$ to $Y$ at a given level of $X_2$ can be obtained from the following expression, which is a simple algebraic manipulation of the conventional moderated regression equation:

$$Y = (b_0 + b_2X_2) + (b_1 + b_3X_2)X_1 + e$$

The simple slope relating $X_1$ to $Y$ at a level of $X_2$ is the compound coefficient $b_1 + b_3X_2$. A t-test can be conducted by dividing this expression by an estimate of its standard error, which is obtained by taking the square root of the variance of $b_1 + b_3X_2$. The variance of this expression can be written as:

$$V(b_1 + b_3X_2) = V(b_1) + X_2^2V(b_3) + 2X_2C(b_1,b_3)$$

Note that, in this equation, $X_2$ is a value that represents a point of interest on the $X_2$ scale. Such points could include the mean of $X_2$, values one standard deviation above and below the mean of $X_2$, or other values chosen for theoretical or practical reasons.

Another method for testing simple slopes is to center the moderator variable at the value at which the simple slope is to be tested, recompute the product term, reestimate the regression equation using the independent variable, the rescaled moderator variable, and the recomputed product term, and test the coefficient on the independent variable. This approach draws from the principle that the coefficient on the independent variable equals the slope of the independent variable when the moderator variable equals zero. Thus, shifting the zero point of the moderator variable to a value that corresponds to the desired simple slope means that the coefficient on the independent variable represents the simple slope of interest. This procedure can be repeated by centering the moderator variable at values that correspond to different simple slopes, recomputing the product term, and reestimating the regression equation.

Simple slopes can also be tested using other procedures for testing linear combinations of regression coefficients. These procedures include matrix approaches in general linear modeling programs and methods that involve imposing the null hypothesis that a simple slope equals zero on a regression equation and testing the reduction in $R^2$ relative to the original moderated regression equation in which this constraint is relaxed.
Scaling Issues:

- In a regression equation using $X_1$, $X_2$, and $X_1X_2$ as predictors, adding or subtracting a constant to $X_1$ changes the coefficient on $X_2$, and vice-versa, provided that the coefficient on $X_1X_2$ is nonzero. Such transformations are permissible for interval measures, which are widely used in the social sciences, because such measures have no true origin (i.e., zero point). The effects of scaling $X_1$ and $X_2$ can be seen with the following variables:

$$X_1^* = X_1 + c$$
$$X_2^* = X_2 + d$$

Using $X_1^*$, $X_2^*$, and their product in a regression equation predicting $Y$ yields:

$$Y = b_0^* + b_1^*X_1^* + b_2^*X_2^* + b_3^*X_1^*X_2^* + e$$

Substitution yields:

$$Y = (b_0^* + b_1^*(X_1 + c) + b_2^*(X_2 + d) + b_3^*(X_1 + c)(X_2 + d) + e$$

$$= (b_0^* + cb_1^* + db_2^* + cdb_3^*) + (b_1^* + db_3^*)X_1 + (b_2^* + cb_3^*)X_2 + b_3^*X_1X_2 + e$$

Comparing the preceding equation to an equation containing $X_1$ and $X_2$ in their original form yields the following equalities:

$$b_0 = b_0^* + cb_1^* + db_2^* + cdb_3^*$$
$$b_1 = b_1^* + db_3^*$$
$$b_2 = b_2^* + cb_3^*$$
$$b_3 = b_3^*$$

Thus, although adding the constants $c$ and $d$ to $X_1$ and $X_2$ does not influence the coefficient on $X_1X_2$ (i.e., $b_3$), it changes the coefficient on $X_1$ by a factor of $db_3^*$ and the coefficient on $X_2$ by a factor of $cb_3^*$ (the change in the intercept, i.e., $cb_1^* + db_2^* + cdb_3^*$, is usually irrelevant in tests of the joint effects of $X_1$ and $X_2$ on $Y$). Because $c$ and $d$ are arbitrary constants, some investigators consider the coefficients on $X_1$ and $X_2$ equally arbitrary, because values of $c$ and $d$ can be selected to make these coefficients take on virtually any value, provided $b_3$ is nonzero.
This apparent dilemma can be resolved by recalling that $b_1$ represents the slope of $Y$ on $X_1$ where $X_2 = 0$. Selecting different values of $d$ shifts the zero point of $X_2$ and effectively selects different points in the distribution of $X_2$ to evaluate the slope of $Y$ on $X_1$. For example, to test the slope of $Y$ on $X_1$ at the mean of $X_2$, $d$ would be set equal to the negative of the mean of $X_2$. This strategy is used to test the slope of $Y$ on $X_1$ at other values of $X_2$, such as one standard deviation above or below the mean of $X_2$. Although the slope of $Y$ on $X_1$ could be forced to take on any arbitrary value based on the selection of $d$, it makes no sense to select values that extend beyond the range of $X_2$, because this would require extrapolating beyond the bounds of the data.

When $X_1$ and $X_2$ are multiplied by arbitrary constants, the coefficients on $X_1$, $X_2$, and $X_1X_2$ will change. However, the standard error of the coefficient on $X_1X_2$ (i.e., $b_3$) will change accordingly, such that tests of significance for $b_3$ are unaffected. The interpretations of $b_1$ and $b_2$ will also remain the same, representing the slope for $X_1$ at $X_2 = 0$ and the slope for $X_2$ for $X_1 = 0$, respectively.

**Multicollinearity:**

- The product $X_1X_2$ is often highly correlated with $X_1$ and $X_2$, as when both $X_1$ and $X_2$ take on only positive values. This phenomenon represents non-essential ill conditioning and can be reduced by rescaling $X_1$ and $X_2$ such that they are centered near zero (this is often accomplished by centering $X_1$ and $X_2$ at their means). Note that rescaling $X_1$ and $X_2$ has no effect on tests of $b_3$, but it changes the interpretation of $b_1$ and $b_2$, such that they represent the slopes of $Y$ on $X_1$ and $X_2$ at the rescaled zero points of $X_2$ and $X_1$, respectively. Rescaling $X_1$ and $X_2$ might facilitate interpretation, but there is little reason to rescale them to reduce multicollinearity, because substantive interpretation is unaffected by the scaling of $X_1$ and $X_2$.

**Reliability:**

- The reliability of a product term is often less than that of either variable constituting the product. If $X_1$ and $X_2$ are orthogonal, the reliability of $X_1X_2$ is the product of the reliabilities of $X_1$ and $X_2$, which is necessarily lower than the reliability of either $X_1$ or $X_2$. If $X_1$ and $X_2$ are positively correlated, the reliability of $X_1X_2$ will be somewhat higher than the product of the reliabilities of $X_1$ and $X_2$. Measurement error drastically reduces the power to detect moderating effects.