

psyc3010 lecture 9

moderated multiple regression

Last week: hierarchical regression
next week: break!
In two weeks: within subjects anova

last week → this week

- Last week we looked at how regression can be performed simultaneously (**standard multiple regression**) or sequentially (**hierarchical multiple regression**)
- This week we cover the most common application of hierarchical multiple regression, namely testing interactions between variables: **moderated multiple regression** and testing for **categorical variables**
- Next week we talk about applications of HMR such as **mediation**, before moving on to within-subjects ANOVA

methods of regression analysis

Standard:

- all predictors entered simultaneously
- each IV evaluated in terms of what it adds to prediction beyond that afforded by all others

Hierarchical:

- predictors entered in pre-specified order, based on logical or theoretical grounds
- Earlier blocks' IVs are credited with any overlapping variance they share with later blocks' IVs

Testing interactions:

Moderated Multiple Regression

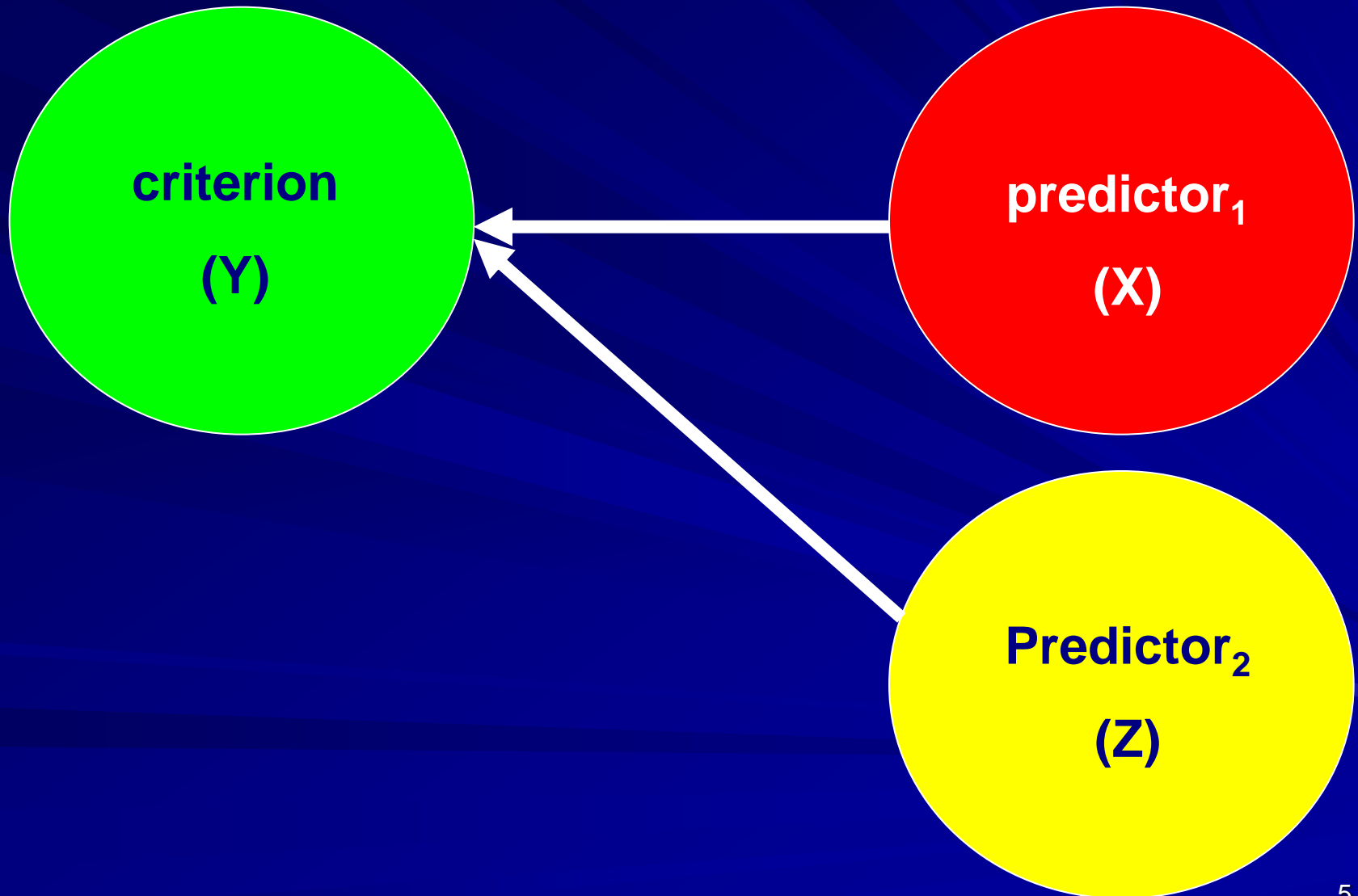
■ what are interactions in MR?:

- the relationship between a criterion and a predictor varies as a function of a second predictor
- the second predictor is usually called a ***moderator***
- moderated regression achieves the same purpose as examination of interactions in factorial anova

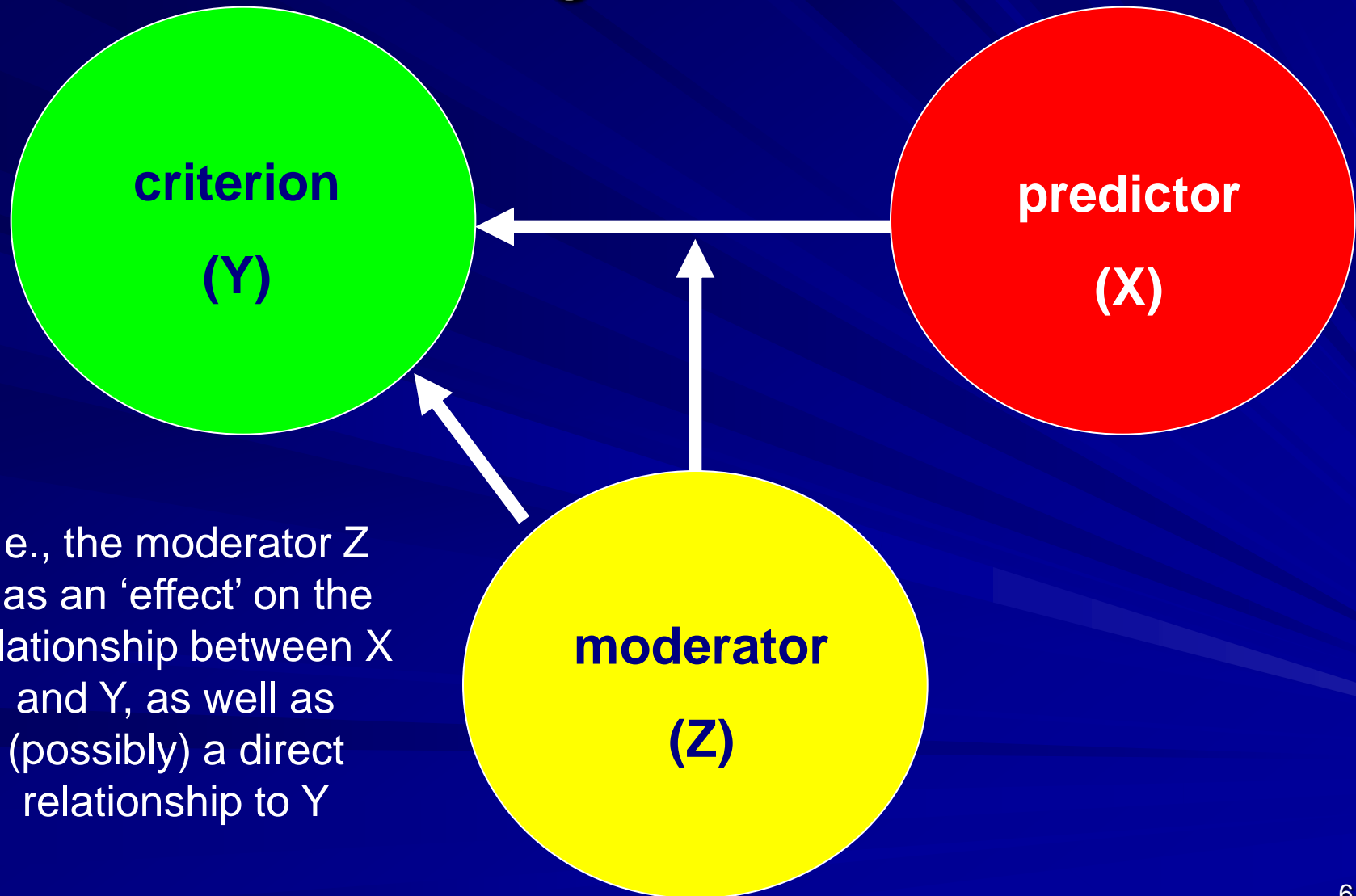
■ examples

- The relationship between stress and health is moderated by social support
- The relationship between IQ and GPA is moderated by motivation
- the relationship between drug use and psychotic episodes *depends upon* ... is enhanced by ... is attenuated by ... a number of dispositional factors (e.g., personality, mental illness in family)

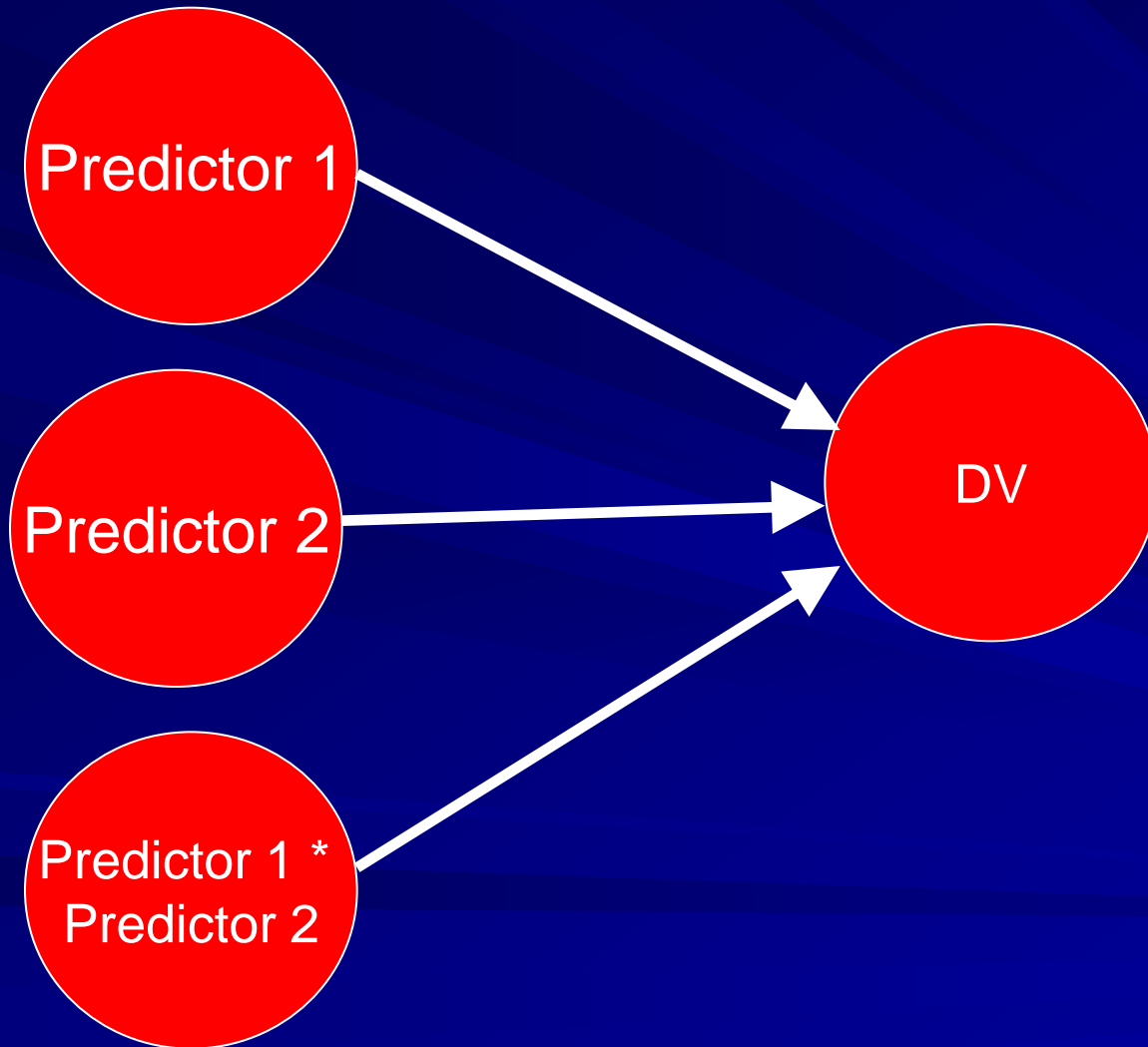
standard regression



Theoretical model of moderated regression



How it's tested



linear model for moderation

- additive effects:

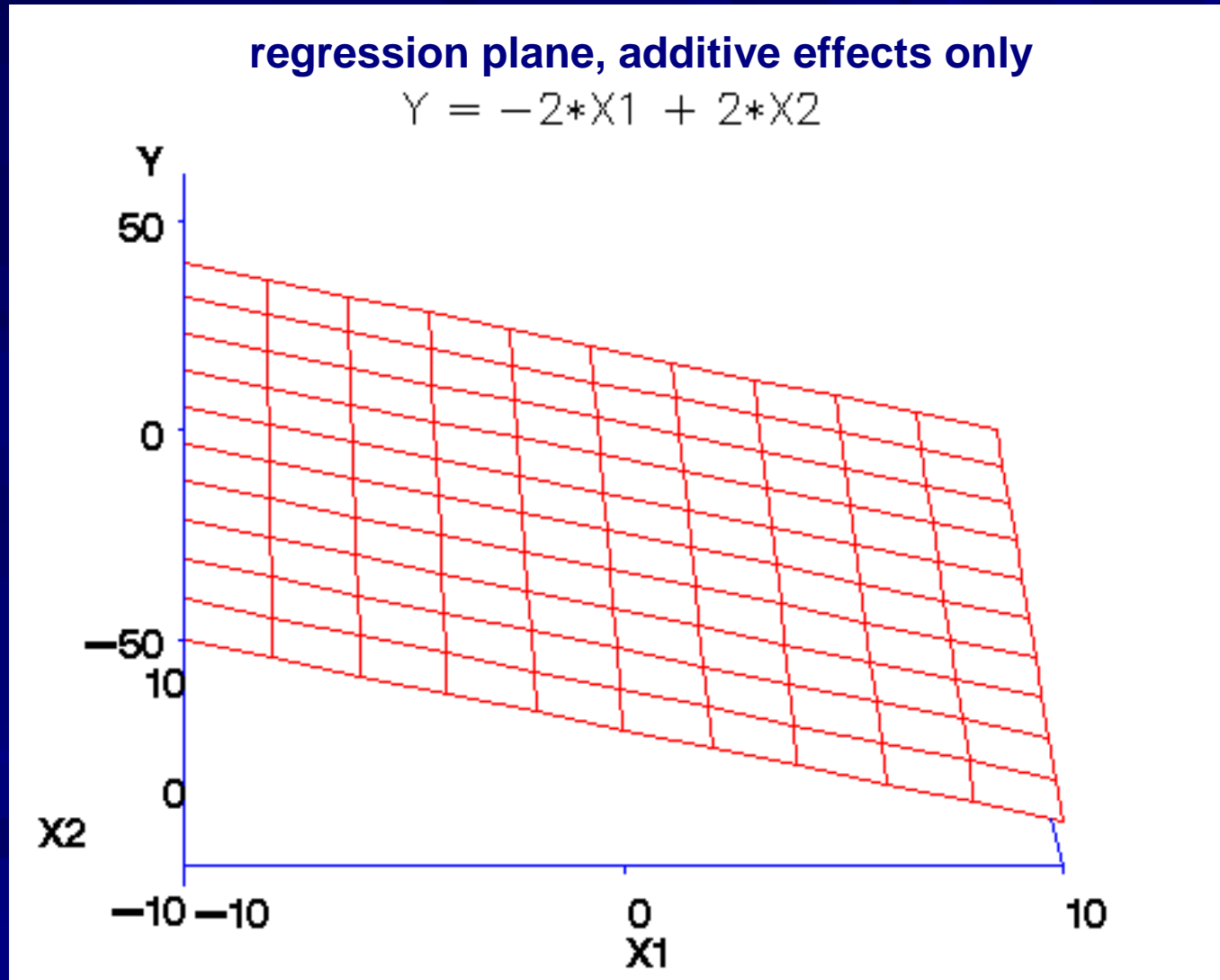
$$\hat{Y} = b_1X + b_2Z + a$$

- ... & non-additive (interactive) effects:

$$\hat{Y} = b_1X + b_2Z + b_3XZ + a$$


Our plan: test whether a model including the interaction term increases the variance accounted for in the DV, compared to a model with just the two IVs considered independently.

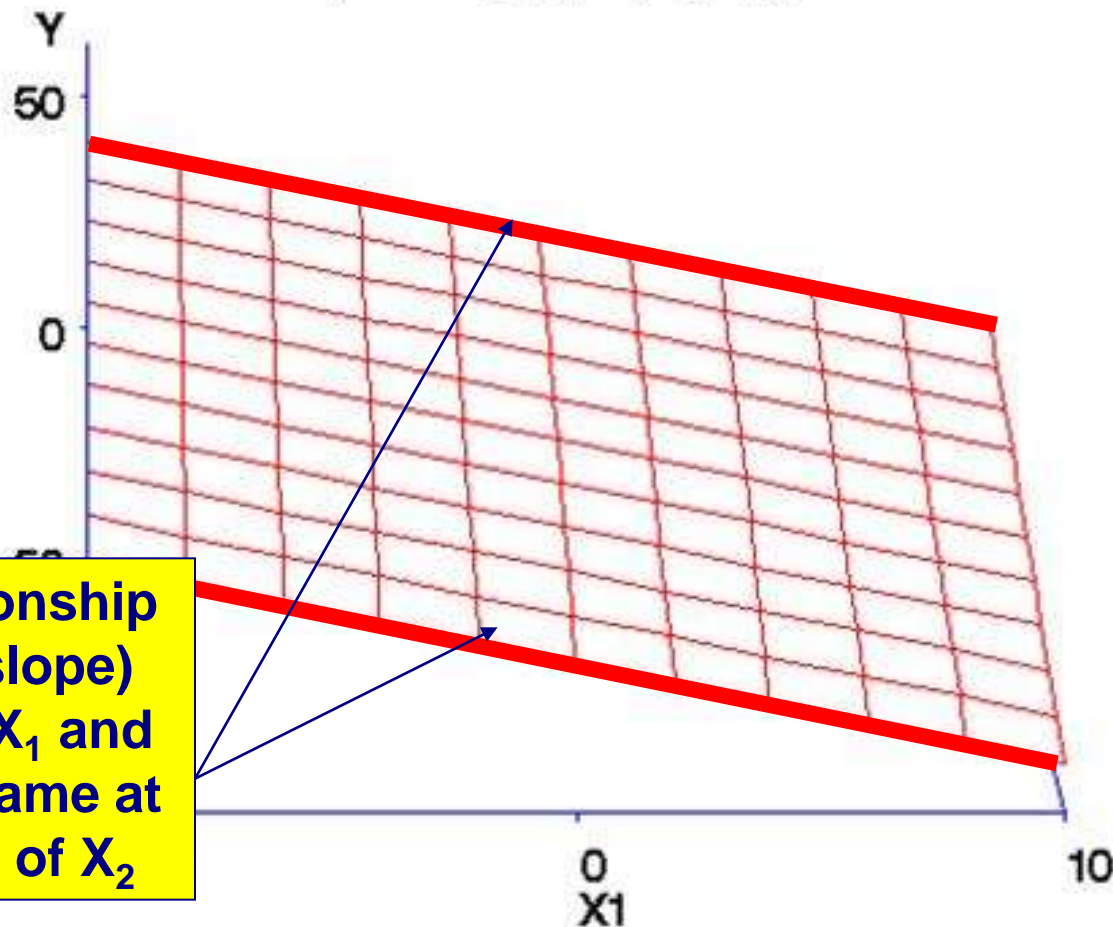
animated graphical representation



animated graphical representation

regression plane, additive effects only

$$Y = -2 \cdot X_1 + 2 \cdot X_2$$

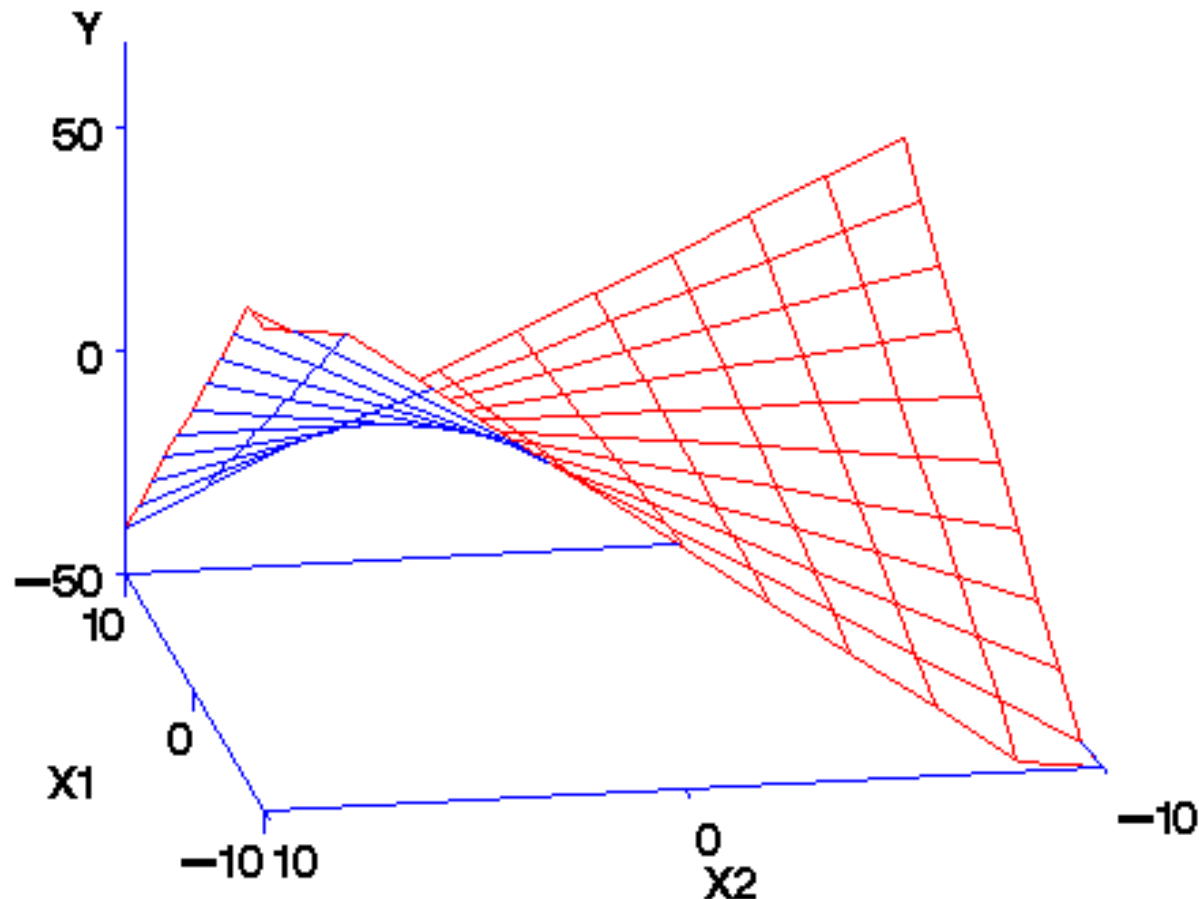


the relationship
(i.e., the slope)
between X_1 and
 Y is the same at
all values of X_2

animated graphical representation

regression plane with interactive effects, varying b_3

$$Y = 0 \cdot X_1 + 1 \cdot X_2 - 0.5 \cdot X_1 \cdot X_2$$

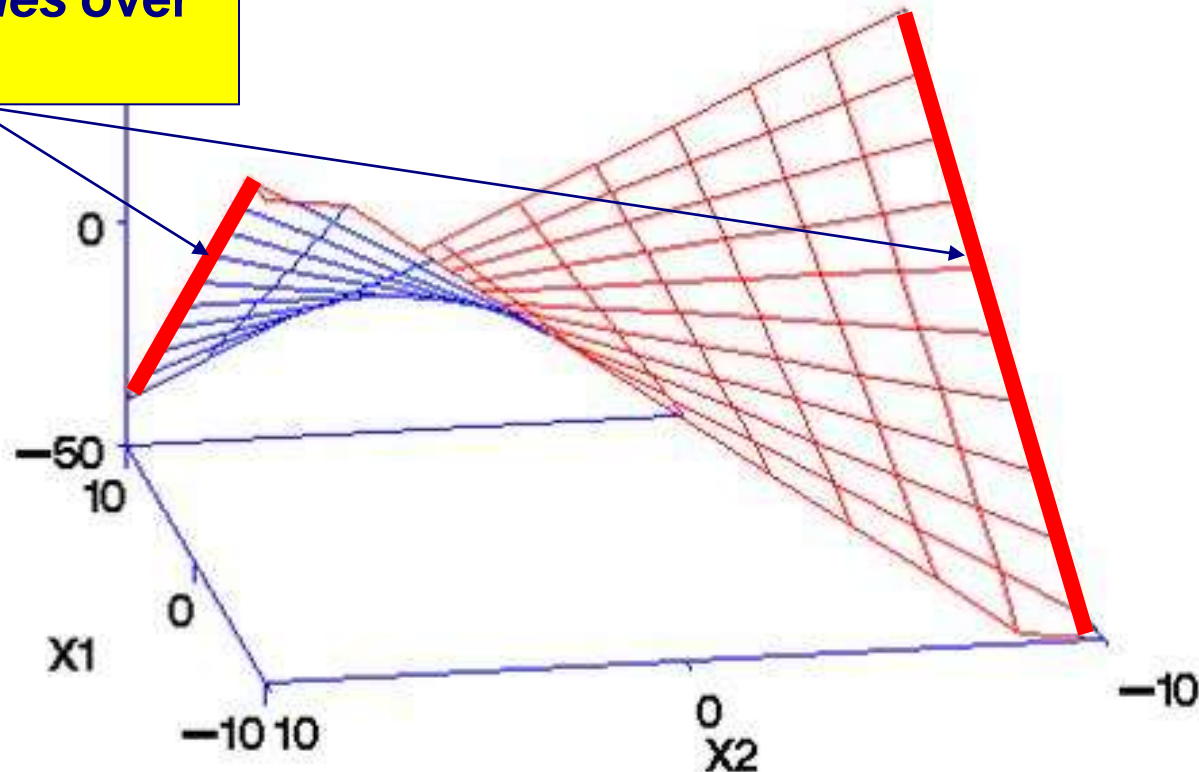


animated graphical representation

regression plane with interactive effects, varying b_3

$$Y = 0 \cdot X_1 + 1 \cdot X_2 - 0.5 \cdot X_1 \cdot X_2$$

the relationship (i.e., the slope) between X_1 and Y varies over values of X_2



moderated regression - 2 issues

1. does the interaction term XZ contribute significantly to prediction of Y ?

- assess via *hierarchical regression* so that the contribution of the interaction term is assessed in a later block after additive effects have been accounted for in first block – sig R^2 increase = sig interaction

2. how do we interpret the effect Z has on the $X \rightarrow Y$ relationship?

- in anova, this was achieved by *simple effects of IV1* at different levels of IV2
- Similarly, in moderated regression, this is achieved by *simple slopes of X* at different values of Z

example

- back to predicting GPA...
 - suppose we know that school performance (OP) and motivation are good predictors of university performance (GPA)
 - but why do many people get a low OP at school and then get a high GPA at university?
- perhaps motivation influences (i.e., *moderates*) the relationship between school performance and university performance

[note – I have *reverse scored* OP for this example so that a higher OP = better school performance]

preliminary statistics

OP and motivation
are both correlated
with GPA

(high validities) 😊

Correlations

OP and motivation
are not correlated
with each other

(low collinearity) 😊

		GPA	OP	Motivation
GPA	Pearson Correlation	1	.319**	.338**
	Sig. (2-tailed)	.	.000	.000
	N	147	147	147
OP	Pearson Correlation	.319**	1	.018
	Sig. (2-tailed)	.000	.	.825
	N	147	147	147
Motivation	Pearson Correlation	.338**	.018	1
	Sig. (2-tailed)	.000	.825	.
	N	147	147	147

** . Correlation is significant at the 0.01 level (2-tailed).

standard regression

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.460 ^a	.212	.201	.54438

a. Predictors: (Constant), Motivation, OP

together, OP and motivation account for a substantial amount of variance in GPA (21%)

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations		
		B	Std. Error	Beta			Zero-order	Partial	Part
1	(Constant)	4.115	.231		17.831	.000			
	OP	.059	.014	.313	4.224	.000	.319	.332	.313
	Motivation	.212	.047	.332	4.485	.000	.338	.350	.332

a. Dependent Variable: GPA

their individual contributions are also substantial and significant

Onward!

The basic steps:

- Calculate interaction term
- Test for significance of interaction
- If interaction is significant, test for simple effects (in this case, simple slopes)

1. Calculating the interaction term

- standard regression deals with the additive effects, but now we just need to compute the interaction term so we can test

$$\hat{Y} = b_1X + b_2Z + b_3XZ + a$$

- Annoyingly SPSS does not do this for us – we have to manually create the interaction term as a new variable, and then add it into the regression equation
- to do this we could simply compute a new variable by multiplying X and Z together
 - **however** this would make the resulting interaction term highly collinear with our other predictors – in other words, the interaction term would be correlated with the original IVs!

calculating the interaction term

(hypothetical data)

participant	OP	motivation	OP x MOT	GPA
	(X)	(Z)	(XZ)	(Y)
1	4	1	4	3.5
2	6	2	12	3.9
3	8	3	24	4.5
4	10	4	40	5.1
5	12	5	60	5.5

the interaction term formed by calculating the cross products of the original predictors will always be correlated with those predictors

mean-centering: a helpful transformation

to mean-center our variables we subtract the mean of each predictor from each observation's score on that predictor:

participant	cOP	cmotivation	cOP x cMOT	GPA
	(X)	(Z)	(XZ)	(Y)
1	-4	-2	8	3.5
2	-2	-1	2	3.9
3	0	0	0	4.5
4	2	1	2	5.1
5	4	2	8	5.5

the interaction term formed by calculating the cross products of the centered predictors will not be (as) correlated with those predictors

Correlations

		INT
OP	Pearson Correlation	.524**
	Sig. (2-tailed)	.000
	N	147
MOT	Pearson Correlation	.839**
	Sig. (2-tailed)	.00
	N	147

**without
centering**

Correlations

		C_INT
C_OP	Pearson Correlation	✓ .014
	Sig. (2-tailed)	.865
	N	147
C_MOT	Pearson Correlation	✓ .105
	Sig. (2-tailed)	.206
	N	147

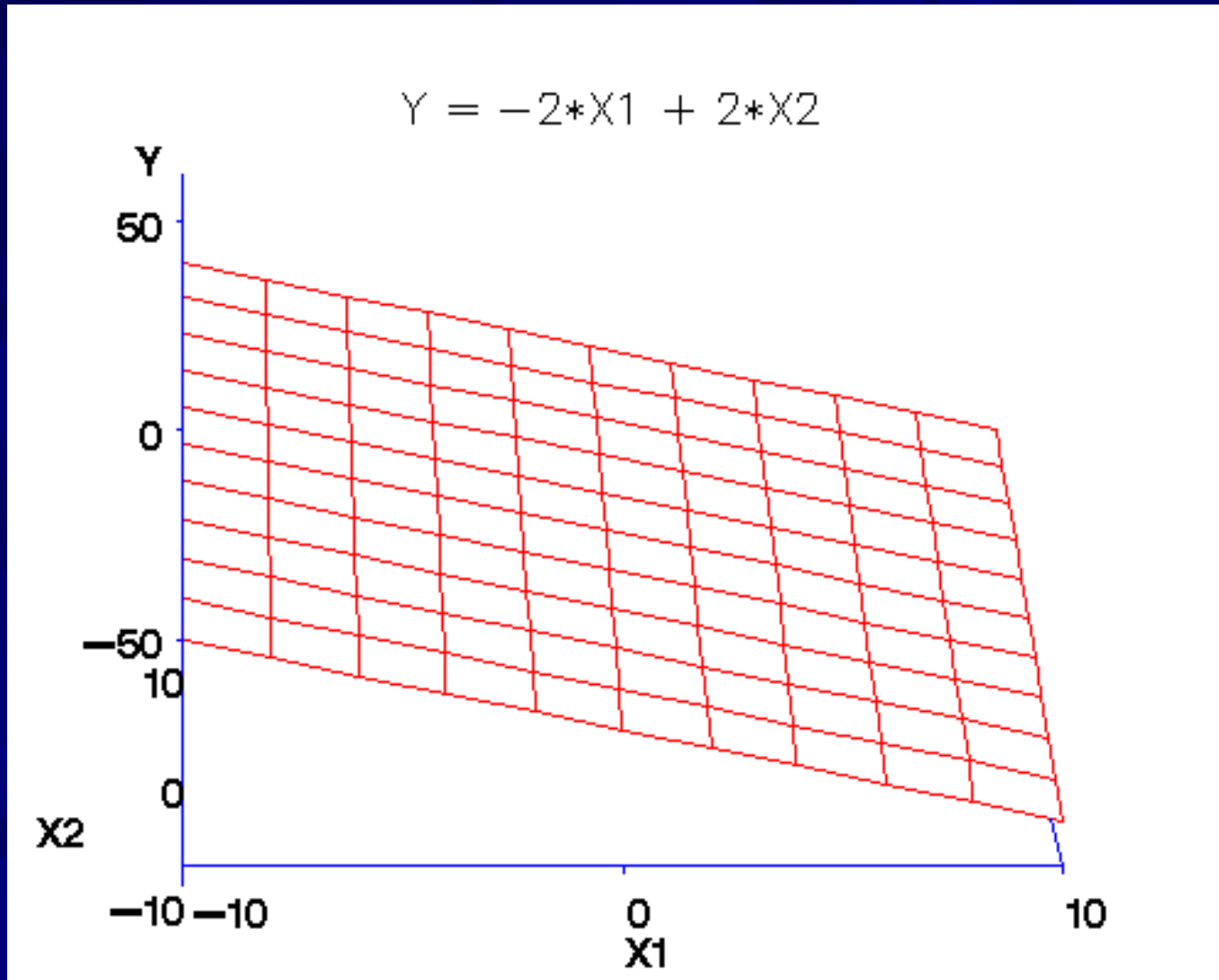
**with
centering**

mean-centering:

an even more helpful transformation

- all we have done is changed the scale so that the mean is now zero – reduces collinearity – but also second advantage...
- technically, b_1 tells us the relationship between a predictor (X) and criterion (Y) ***when all the other predictors (e.g., Z) are zero***
- in standard regression this doesn't matter, because the relationship between X and Y is ***the same*** at all values of Z (i.e., there are no interactions)
 - so the slope of X in relation to Y is the same when Z is zero as at all other values of Z – test is meaningful

regression of Y on X (i.e., X_1) is the same at all values of Z (i.e., X_2)



- but if there is a XZ interaction, then the relationship between X and Y ***changes across values of Z***
- Then we must ask, is 'Z=zero' the most representative level at which to consider the relationship of X and Y (test coefficient of X)?
 - What if Z ranges from 1 to 7 ? Then $z=0$ doesn't even exist in the data!
- By using centered predictors, b_1 represents the relationship between X and Y at the mean of Z
- Point: We always use mean centering to calculate interaction terms
 - to reduce collinearity in moderated multiple regression
 - to increase ease of interpretation of coefficients by making 'the slope when all other predictors are at zero' meaningful [zero=mean]

2. Testing the interaction term for significance

- once we have centered our predictors and computed an interaction term (a new variable which is the cross product of the two centered predictors – $cX_1X_2 = cX_1 * cX_2$) we can test the interaction for significance
- hierarchical regression is used
- at step 1 we enter centered IVs (OP and Motivation) as predictors of DV (GPA). These are the additive effects.
- at step 2 we enter the interaction term (previously calculated as product of two centered predictors) to see if this accounts for additional variance in GPA

preliminary statistics

our new centered predictors were computed by subtracting their mean from all values – hence their new mean is zero

(recall OP is reverse scaled!)

standard deviations have not changed

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
GPA	147	3.40	6.50	5.4754	.60891
C_OP	147	-12.92	5.08	.0004	3.24039
C_MOT	147	-1.56	2.44	.0000	.95284
Valid N (listwise)	147				

note that we have not centered the criterion (DV) – this is simply because there is no need (i.e., we don't have the collinearity problem) and it makes sense to interpret the solution in terms of the scale GPA is on

preliminary statistics

see how we haven't influenced any of the original correlations by centering

also, interaction term is correlated with GPA

Correlations

		GPA	C_OP	C_MOT	C_INT
GPA	Pearson Correlation	1	.319**	.338**	-.204*
	Sig. (2-tailed)	.	.000	.000	.013
	N	147	147	147	147
C_OP	Pearson Correlation	.319**	1	.018	.014
	Sig. (2-tailed)	.000	.	.825	.865
	N	147	147	147	147
C_MOT	Pearson Correlation	.338**	.018	1	.105
	Sig. (2-tailed)	.000	.825	.	.206
	N	147	147	147	147
C_INT	Pearson Correlation	-.204*	.014	.105	1
	Sig. (2-tailed)	.013	.865	.206	.
	N	147	147	147	147

** . Correlation is significant at the 0.01 level (2-tailed).

* . Correlation is significant at the 0.05 level (2-tailed).

moderated regression

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df 1	df 2	Sig. F Change
1	.460 ^a	.212	.201	.54438	.212	19.335	2	144	.000
2	.521 ^b	.271	.256	.52520	.060	11.706	1	143	.001

a. Predictors: (Constant), C_MOT, C_OP

b. Predictors: (Constant), C_MOT, C_OP, C_INT

at step 1, results are *identical* to the standard regression with OP and motivation as predictors


at step 2, $R^2_{ch} = .06$, so the interaction explains 6% of the variance in GPA over and above the additive effects of motivation and OP
furthermore, this increment in explained variance is significant,
 $F_{ch}(1,143) = 11.71, p = .001$

moderated regression

- here we have addressed the first issue in moderated regression and state that there is a significant interaction between OP and Motivation in the prediction of GPA
- next step is to follow this interaction up similarly to how we would in anova
 - analysis of ***simple slopes***

3. simple slopes

They have levels e.g.
1-7 – who wants to
do seven simple
slopes? Ewww!

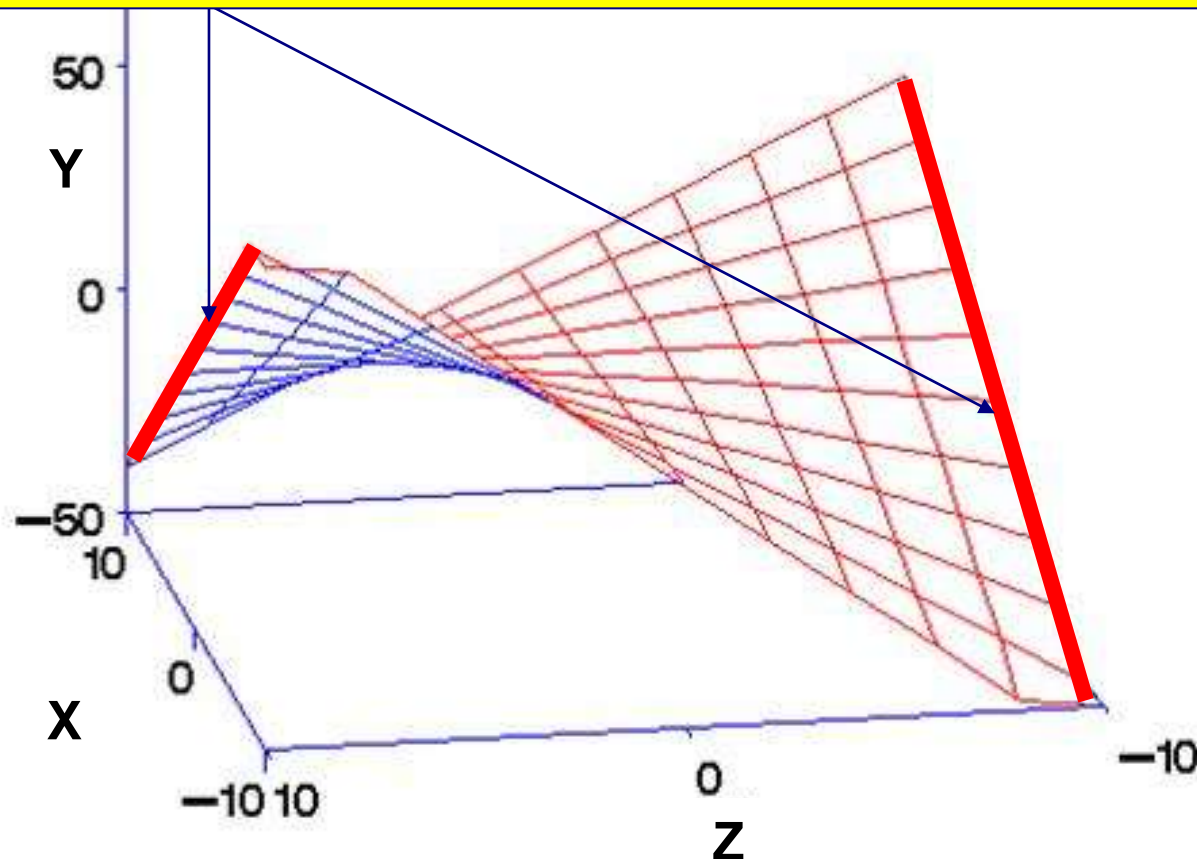


- in anova, simple effects examine the effect of one factor at each level of the other factor
- in moderated regression, our predictors are continuous – i.e., they don't have levels
- we select critical values of the moderator where it is interesting to examine the simple slopes of the IV in relation to Y
- We use logical grounds, usually +1 and -1 SD of moderator (“high” and “low” levels of Z)

simple slopes

we examine the relationship between X and Y at high and low values of Z

(& typically the values chosen are ± 1 SD of Z)



back to the regression output...

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations
		B	Std. Error	Beta			Part
1	(Constant)	5.475	.045		121.949	.000	
	C_OP	.059	.014	.313	4.224	.000	.313
	C_MOT	.212	.047	.332	4.485	.000	.332
2	(Constant)	5.478	.043		126.442	.000	
	C_OP	.059	.013	.316	4.420	.000	.315
	C_MOT	.229	.046	.358	4.982	.000	.356
	C_INT	-.047	.014	-.246	-3.421	.001	-.244

a. Dependent Variable: GPA

at step 2, we get the partial regression coefficients and so can build our regression equation specifying additive and interactive effects:

$$\hat{Y} = 0.059C_OP + 0.229C_MOT + -0.047C_INT + 5.478$$

Hand calcs with the linear model

as we are conceptualising the simple slopes of X at Z1 and Z2 as two slopes of Y regressed upon X at high and low levels of Z, we can employ the equation accordingly, i.e., take the original unstandardised regression equation

$$\hat{Y} = B_1X + B_2Z + B_3XZ + c$$

For low Z (-1SD), replace Z with -# (where # = the SD of Z)

For high Z (+1SD), replace Z with +# (where # = the SD of Z)

Deriving the simple slopes by hand

So for the simple slopes for OP at high and low levels of motivation:

$$\hat{Y} = 0.059C_OP + 0.229C_MOT + -0.047C_OP \times MOT + 5.478$$

1) at high motivation (where $C_MOT = 0.9528$), it becomes:

$$\begin{aligned}\hat{Y} &= 0.059C_OP + 0.229(0.9528) + -0.047(C_OP \times 0.9528) + 5.478 \\ &= (0.059 - 0.0448)C_OP + (0.2182 + 5.478) \\ &= 0.0142C_OP + 5.6962\end{aligned}$$

I.e., the simple slope of OP at high motivation is .014

deriving the simple slopes

and, at low motivation (where C_MOT = - 0.9528)

$$\begin{aligned}\hat{Y} &= 0.059C_OP + 0.229(-0.9528) + -0.047(C_OP \times -0.9528) + 5.478 \\ &= (0.059 + 0.0448)C_OP + (-0.2182 + 5.478) \\ &= 0.1038C_OP + 5.2598\end{aligned}$$

I.e., the simple slope of OP at low motivation is .104

plotting the simple slopes

we can now use these equations to plot our interaction

– each specifies the relationship between OP and GPA at selected values of motivation (e.g., high and low)

- if we sub in 2 levels of OP for each we will get 4 points to plot on a graph...

so, for high OP (SD = 3.2404, so 1SD above the mean of OP = +3.2404)

for low motivation: $\hat{Y} = 0.1038 \times 3.2404 + 5.2598 = \underline{5.5962}$

for high motivation: $\hat{Y} = 0.0142 \times 3.2404 + 5.6962 = \underline{5.7422}$

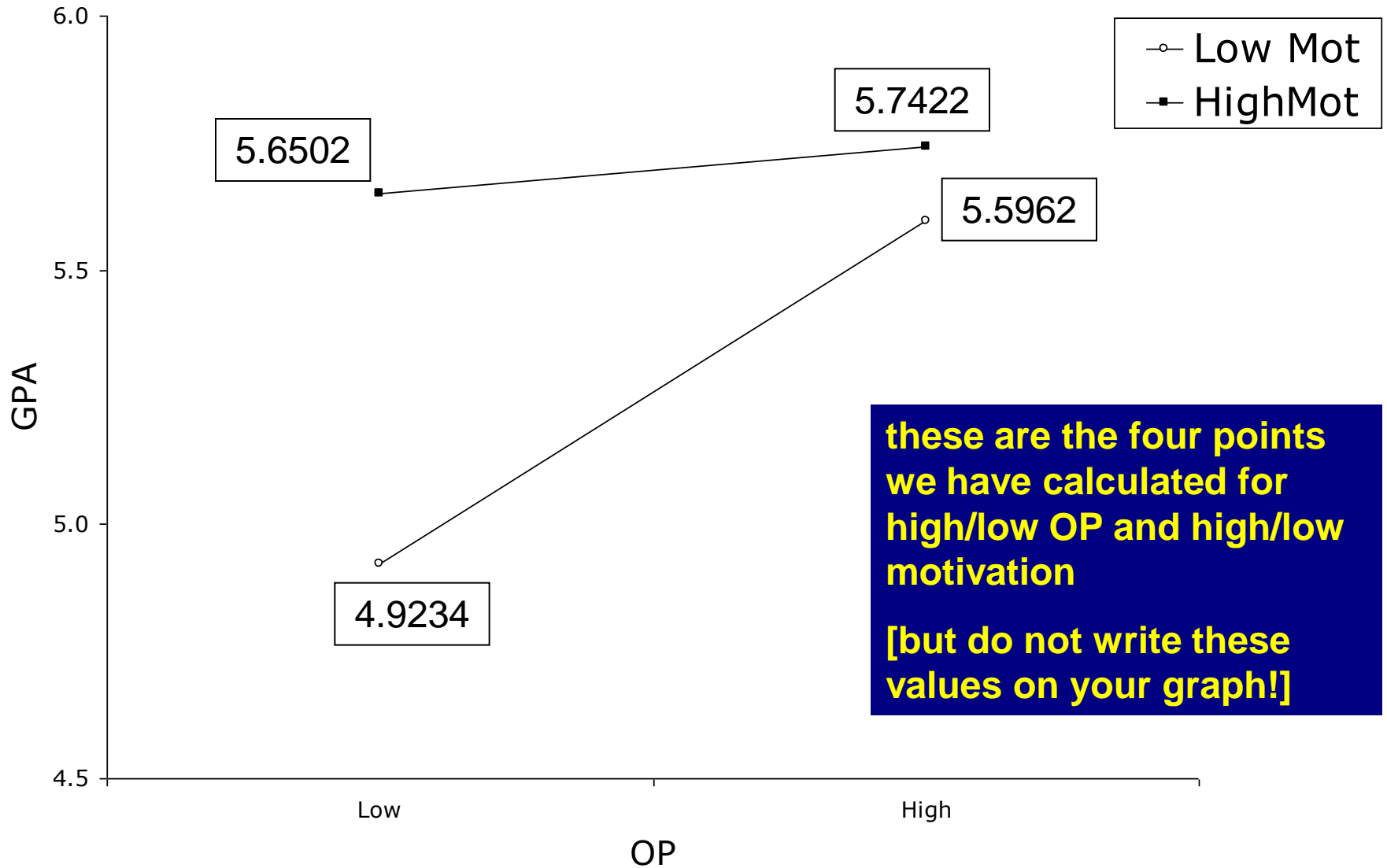
plotting the simple slopes

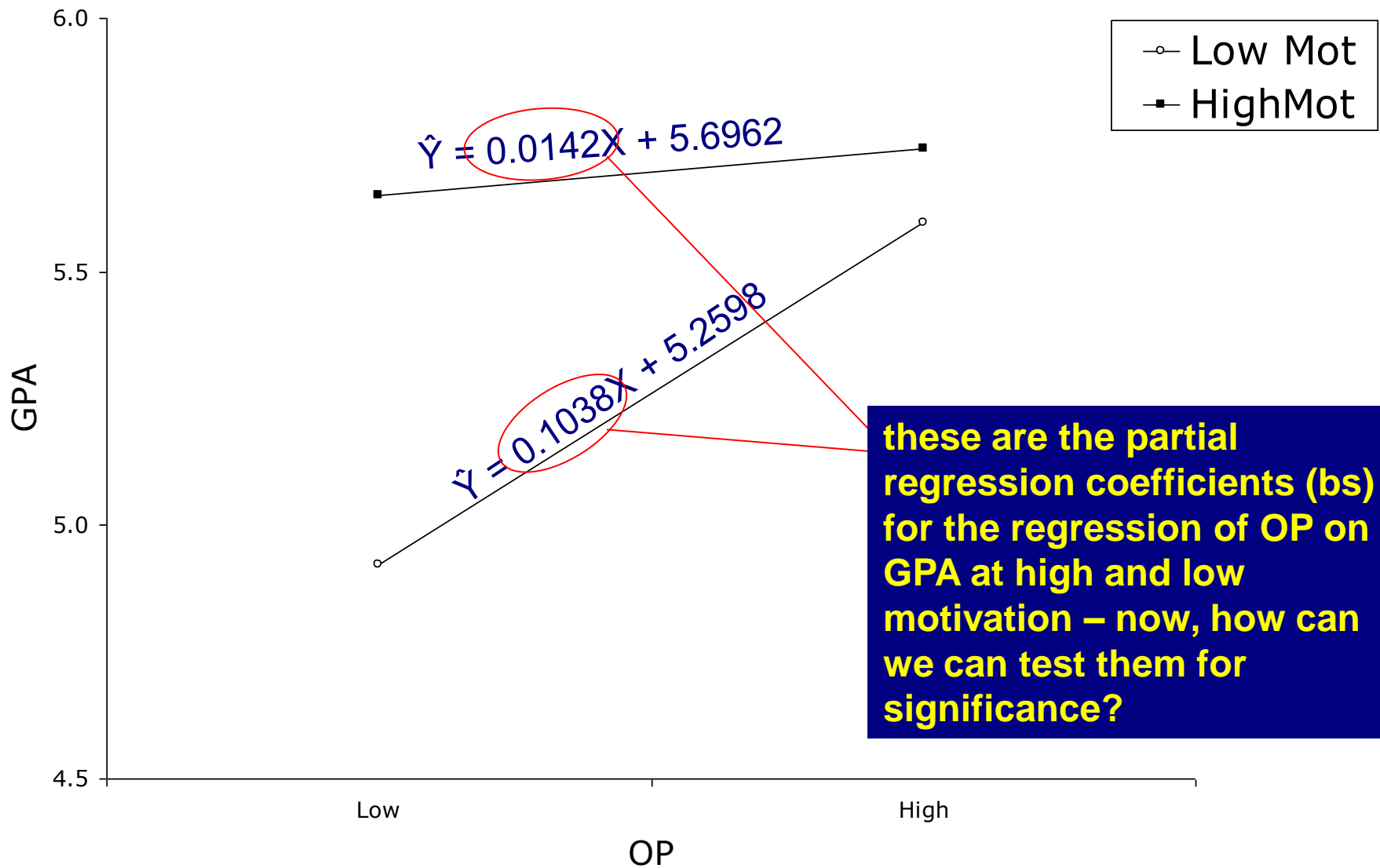
we can now use these equations to plot our interaction – each specifies the relationship between OP and GPA at selected values of motivation, if we sub in 2 levels of OP for each we will get 4 points to plot on a graph...

and, for low OP (1SD below the mean of OP = - 3.2404)

for low motivation: $\hat{Y} = 0.1038 \times -3.2404 + 5.2598 = \underline{4.9234}$

for high motivation: $\hat{Y} = 0.0142 \times -3.2404 + 5.6962 = \underline{5.6502}$





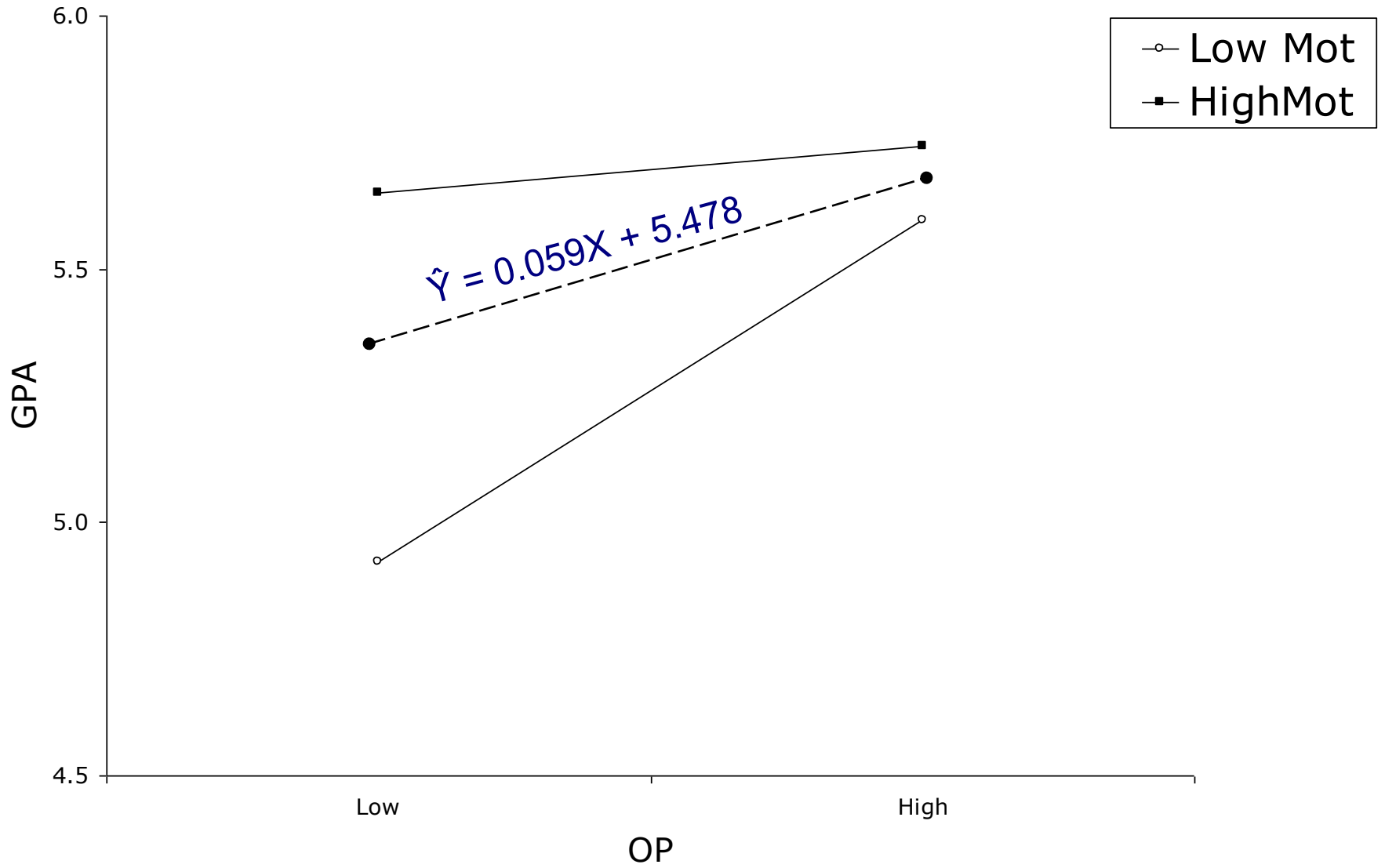
Testing the significance of the simple slopes

- remember that b_1 is the slope for $X \rightarrow Y$ when $Z = 0$?
- since we have *centered* our predictors, 0 is the mean of Z , hence default test of b_1 is the slope for $c_op \rightarrow Y$ at the mean value of Z

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations
		B	Std. Error	Beta			Part
1	(Constant)	5.475	.045		121.949	.000	
	C_OP	.059	.014	.313	4.224	.000	.313
	C_MOT	.212	.047	.332	4.485	.000	.332
2	(Constant)	5.478	.043		126.442	.000	
	C_OP	.059	.013	.316	4.420	.000	.315
	C_MOT	.229	.046	.358	4.982	.000	.356
	C_INT	-.047	.014	-.246	-3.421	.001	-.244

a. Dependent Variable: GPA



Testing the significance of the simple slopes

- remember that b_1 is the slope for $X \rightarrow Y$ when $Z = 0$?
- since we have *centered* our predictors, 0 is the mean of Z , hence default test of b_1 is the slope for $c_op \rightarrow Y$ at the mean value of Z

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients			Correlations
		B	Std. Error				
1	(Constant)	5.475	.045				
	C_OP	.059	.014				
	C_MOT	.212	.047				
2	(Constant)	5.478	.043				
	C_OP	.059	.013				
	C_MOT	.229	.046				
	C_INT	-.047	.014				
				-.246	-3.421	.001	-.244

How can we make SPSS test the slope of OP when mot = +1 SD?

logically, if we now center Z at +1 SD, then b_1 would be the slope for $X \rightarrow Y$ at high Z

a. Dependent Variable: GPA

To test the simple slopes for high & low moderator:

Step 1. Create two variables, one for each level of the moderator you are interested in, using the formula $Mod_{CV} = CMod - CV$.

Critical value is normally 1 SD (Standard deviation)

$$Mod_{ABOVE} = cmod - (SD)$$

$$Mod_{BELOW} = cmod - (-SD)$$

Be consistent in your labelling to help remember what each variable is – include cues to remind you it's centered at + or -1SD.

$$\text{e.g., } c_mothi = cmot - (.95)$$

$$\text{And } c_motlo = cmot - (-.95).$$

$$\text{Or e.g., } highc_mot = cmot - (.95)$$

$$\text{And } lowc_mot = cmot - (-.95).$$

To test the simple slopes for high & low moderator:

Step 2. Form the crossproducts of the new variable with the other predictor IV , by computing $IV * Mod_{CV}$

e.g., I like to use the prefix `ci_` to stand for centered interaction.

`compute ci_opxmhi = c_op * c_mothi .`

`compute ci_opxmlo = c_op * c_motlo .`

`execute`

As long as you're consistent in your naming so you remember which is which, you'll be right.

To test the simple slopes for high & low moderator:

Step 3. Regress Y on IV , Mod_{CV} , & $IV*Mod_{CV}$ for each of the 2 critical values (high and low).

- Examine test of IV in final block with $MODCV$ and interaction of IV & $MODCV$ also in the equation.
- I.e., test b_{IV} for the simple slope of Y on IV at the conditional value of Mod

Test simple slope of OP at high motivation:

- first, center motivation at **+1SD** and re-run the MMR, i.e.,
 - subtract **0.9528** from C_MOT to give **C_MOThigh**
 - recalculate interaction term (C_OP x C_MOThigh) to give **INT_HI**
 - enter C_OP and C_MOThigh at step 1 and INThigh at step2

Model		Unstandardized Coefficients		Standardized Coefficients			
		B	Std. Error				
1	(Constant)	5.678	.064				
	C_OP	.059	.014				
	MOTHIGH	.212	.047	.552	4.465	.000	.552
2	(Constant)	5.696	.062		92.458	.000	
	C_OP	.014	.019	.076	.764	.446	.055
	MOTHIGH	.229	.046	.358	4.982	.000	.356
	C_INT_HI	-.047	.014	-.341	-3.421	.001	-.244

b_1 is now the slope for $X \rightarrow Y$ at high Z – identical to the one we have calculated by hand, and now we have the SE value and t-test.

a. Dependent Variable: GPA

Test simple slope of OP at low motivation:

- second, center motivation at **-1SD** and re-run the MMR, i.e.,
 - subtract **-0.9528** from C_MOT to give **C_MOT-**
 - recalculate interaction term (C_OP x C_MOT-) to give **INT-**
 - enter C_OP and C_MOT- at step 1 and INT- at step2

Model		Unstandardized Coefficients		Standardized Coefficients			
		B	Std. Error				
1	(Constant)	5.273	.064				
	C_OP	.059	.014				
	MOTLOW	.212	.047	.332	4.485	.000	.332
2	(Constant)	5.260	.061		85.551	.000	
	C_OP	.104	.019	.555	5.517	.000	.394
	MOTLOW	.229	.046	.358	4.982	.000	.356
	C_INT_LO	-.047	.014	-.346	-3.421	.001	-.244

b_1 is now the slope for $X \rightarrow Y$ at low Z – identical to the one we have calculated by hand, and now we have the SE value! (and the t-test!)

a. Dependent Variable: GPA

phew...

- so now we have addressed the second issue in moderated regression – the relationship between OP and GPA is significant at lower levels of motivation but not higher levels
 - High (current) motivation weakens the effects of prior academic performance on current academic performance
 - More examples in the tutes

Review: The basic steps

- 1. Calculate interaction term
 - Center each of the independent variables
 - Multiply the centered variables to create a product term $int = civ1 * civ2$
- 2. Test for significance of interaction
 - Conduct a HMR with $civ1$ and $civ2$ in $B1$, and the centered interaction term in $B2$
 - A sig R^2 ch / interaction coefficient in $B2$ means the interaction is significant
- 3. If interaction is significant, test for simple effects (in this case, simple slopes):
 - Choose critical values of moderator that are interested in – usually ($\pm 1SD$)
 - Calculate two simple regression equations for each critical value (usually one for participants 1 SD above mean on moderator, one for those 1 SD below mean; could also do more but 2 is standard)
 - plot the interaction (one simple regression line of the IV for each critical value of the moderator that we are interested in)
 - test significance of slopes of simple regression lines in SPSS
 - Report significance test and effect size (squared semi-partial r) for each simple slope

limitations of MMR

- power – it is very difficult to detect statistical interactions in MR
 - 147 cases for this example where usually the examples have only 20 or so people
- effect sizes of the interaction are usually small (around 5%)
 - some discussion of this issue is provided by Jaccardi & Wan, 1995
- Compared to ANOVA, much more difficult to implement in normal statistical packages – no dedicated program.

some extra issues – predictor importance

- note that in the Z_{high} and Z_{low} regressions we have just run, we get the sr^2 values for our simple slopes:
 - for C_OP in the Z_{high} solution, $sr = 0.055$ (reported as part r)
OP accounts for .3% of the variance at high levels of motivation
 - for C_OP in the Z_{low} solution, $sr = 0.394$
OP accounts for 15.5% of the variance at low levels of motivation
- standardised partial regression coefficients might also be useful for interpretation if we aren't familiar with the scale on the criterion
 - e.g., currently we know that, *at low levels of motivation, a 1 unit increase in OP results in a .104 unit increase in GPA*
 - not very informative if you are reading this paper in Europe/America!

Can we just use the β values provided by SPSS in MMR?

- the β values for the simple slopes will be correct...but ***the interaction term will be incorrect***
 - because SPSS ***first*** computes the cross product of raw X and Z (we did this manually) and ***then*** standardises that value
 - we want the cross product of standardised X and Z, not the standardised cross product of X and Z
- Much simpler to use unstandardised solution – this is perfectly ok, especially where interested in unit changes in the DV (meaningful scale)
 - Point: Sometimes report unstandardised coefficients in MMR – scale dependent, so report sr^2 to get a sense of what ‘size’ the effect is
 - Might rarely perform MMR on standardised scores (too much PITA for most)
- Common alternative is to report interaction as beta but not interpret its magnitude
- Simple slopes can be reported as betas (interpretable)

MMR - additional points

- it is possible to have 3+ predictors (3-way moderated regression, 4-way ...)
 - Extension of 2-way design. (a) centre all variables (b) compute two-way interaction terms ($clV1 * clV2$, etc.) (c) compute 3-way interaction(s) (d) perform HMR with $clVs$ in Block 1, 2-way interactions in Block 2, and 3-way in Block 3. (d) Where necessary, compute simple slopes by substituting in CVs .
- Aiken and West (1991) have written a good book on interactions in MR and this is in the high use section of the library

categorical predictors in hierarchical multiple regression

dummy variables:

- coded 1 to indicate presence of an attribute; 0 = absence
- if variable has J categories, need $J - 1$ dummy variables as predictors in regression

e.g., gender (male/female) can be coded with 1 dummy variable
female = 1, male = 0

e.g., religion

- for ANOVA could code IV 1 = Protestant; 2 = Catholic; 3 = Jewish; 4 = None but in regression makes no sense (no “real” order)

- **Create $j-1$ dummy variables:**

Three
different new
variables

	P	C	J	N
DUV1	1	0	0	0
DUV2	0	1	0	0
DUV3	0	0	1	0

Enter dummy variables together as a block in hierarchical analysis.

- Test for model R^2 change for that block tells you overall effect of religion
- Each coefficient is a follow-up test whose significance tells you if that group differs from baseline group (“no religion” = 0 0 0)

testing the effects of categorical predictors

Effect codes:

- if variable has J categories, need $J - 1$ effect codes as predictors in multiple regression
- With 2 groups, use 1, -1 (e.g., gender: men = -1, women = 1)
- With > 2 groups, can create variables to represent the linear contrasts of interest (need to be orthogonal though!)

e.g., religion 1 = Protestant; 2 = Catholic; 3 = Jewish; 4 = None

	P	C	J	N
EC1	1	1	1	-3
EC2	1	1	-2	0
EC3	1	-1	0	0

Enter all effect codes together as a block in hierarchical analysis.

- Test for model R^2 change - tells you overall effect of IV
- Significance of each coefficient tells you if contrast is significant – e.g., EC1: Do religious differ from non-religious?
- Effect codes good because don't enforce 'baseline' category; allow more complex comparisons (vs dummy codes only pairwise); more robust test of interactions

What about categorical IVs and interactions?

- E.g. if Z is dichotomous (2 levels – e.g., gender, semester):
- Using dummy codes:
 - assign value of 0 & 1 [*not centered!]
 - calculate interaction term using $DUV * \text{centered continuous IV}$; test in HMR
 - plot and test simple regression equations for each group (men, women) with variables centred at 0, 1 instead of +1SD, -1SD
 - Older technique, most common. Problem of multicollinearity and non-orthogonality.
- Can also use “effect” codes:
 - assign value of -1 & 1 [more like centered var.]
 - calculate interaction term using $EC * IV$; test in HMR
 - plot and test simple regression equations for each group (men, women) with variables centered at -1, 1
 - Newer technique – some ease of interpretation – and more appropriate if there are >1 interaction to be tested with categorical variables (b/c < multicollinearity, e.g.)

Categorical IVs as moderators

- More complex if >2 levels
 - need $j-1$ dummy values or effect codes
 - Create $j-1$ interaction terms (centered IV x each dummy variable or effect code)
 - Run HMR with centered IV and original DUVs/ECs in Block 1, Enter interaction terms jointly in block 2
 - Sig of R^2 change in Block 2 = sig of interaction overall
 - Each individual coefficient is a follow-up test
- Interactions between two categorical variables (e.g., gender X employment status) can be tested in regression – in which case the regression solution is *identical* to anova (with equal n)

more advanced topics in MR

- **You need to know all about 2-way MMR in 3010**, but in this course you won't be assessed on categorical predictors or 3-way MMR. You'll see them all the time in journal articles though!
- **For more materials on categorical predictors (SMR or HMR)**
 - see Field (2006): pages 208 → 212
- **3+ predictors in MMR**
 - see Aiken & West (1991)
- **categorical moderator variables**
 - see Aiken & West (1991)

In class next week:

- Mediation
- Within-subjects ANOVA

In the tutes:

- This week: Moderated regression cont'd, SPSS, A2
- Next week: Within-subjects designs

readings :

- For MMR: Howell Ch. 15, pp. 577-582
- For next week: Howell Ch. 14

Note:

- There are additional resources on mediation and moderation on my web site, <http://www2.psy.uq.edu.au/~wlouis/> - they were written for postgrads so may be hard to follow but are included in case they can help.