Everything You Wanted to Know about Moderation
(but were afraid to ask)

Jeremy F. Dawson
University of Sheffield

Andreas W. Richter
University of Cambridge

the Power of Words
74th Annual Meeting of the Academy of Management
August 1-5, 2014
Philadelphia, PA

Resources for this PDW

- Slides
- SPSS data set
- SPSS syntax file
- Excel templates

- Available at

  http://www.jeremydawson.com/pdw.htm

J Bus Psychol (2014) 29:1–19
DOI 10.1007/s10869-013-9308-7

Moderation in Management Research: What, Why, When, and How

Jeremy F. Dawson

Published online: 9 May 2013

Abstract Many theories in management, psychology, and other disciplines rely on moderating variables: those which affect the strength or nature of the relationship between two other variables. Despite the near-ubiquitous nature of such effects, the methods for testing and interpreting them are not always well understood. This article introduces the concept of moderation and describes how moderator effects are tested and interpreted for a series of model types, beginning with straightforward two-way interactions with Normal outcomes, moving to three-way combining structural equation modeling with meta-analysis (c.f., Johnson et al. 2011; Landis 2013). The present article is designed to complement these valuable articles by explaining many of the issues surrounding one of the most common types of statistical model found in the management and organizational literature: moderation, or interaction effects.

Life is rarely straightforward. We may believe that exercising will help us to lose weight, or that earning more money will enable us to be happier, but these effects are

Everything You Wanted to Know about Moderation

- **Many theories are concerned with whether, or to which extent, the effect of an independent variable on a dependent variable depends on another, so called 'moderator' variable**

Everything You Wanted to Know about Moderation

- **Examples:**
  - **Hoever et al. (2012, JAP): The relationship between team diversity and team creativity depends on the level of perspective taking.**
  - **Baer (2012, AMJ): The relationship between the generation of ideas and their implementation depends on both employees' motivation and their ability to network.**

Session organizer

1. **Testing and probing two-way and three-way interactions using MRA**
2. **Curvilinear interactions**
3. **Interactions with non-Normal outcomes**
4. **Extensions of MRA**

Testing two-way interactions

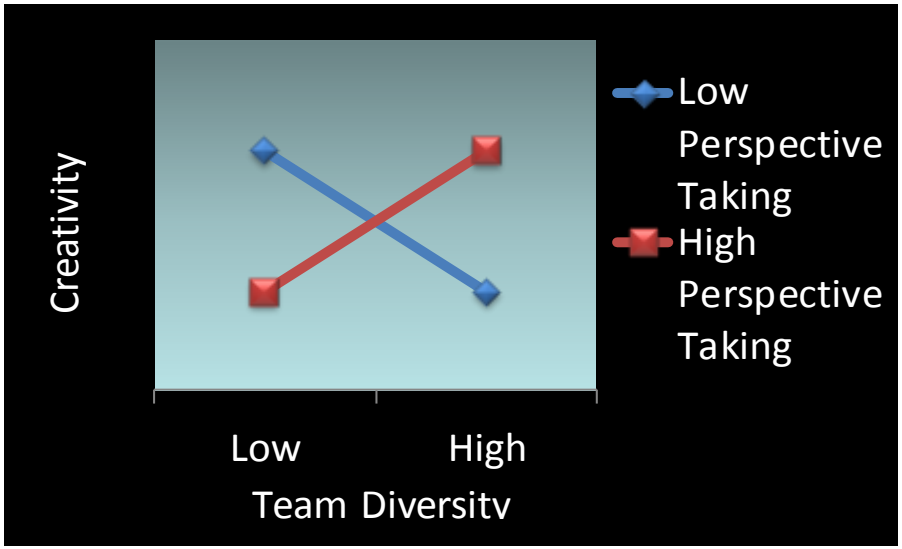
- $\hat{Y} = b_0 + b_1X + b_2Z + b_3XZ$

Predicted *Y*	Intercept	First order effects	Interaction term

Probing two-way interactions

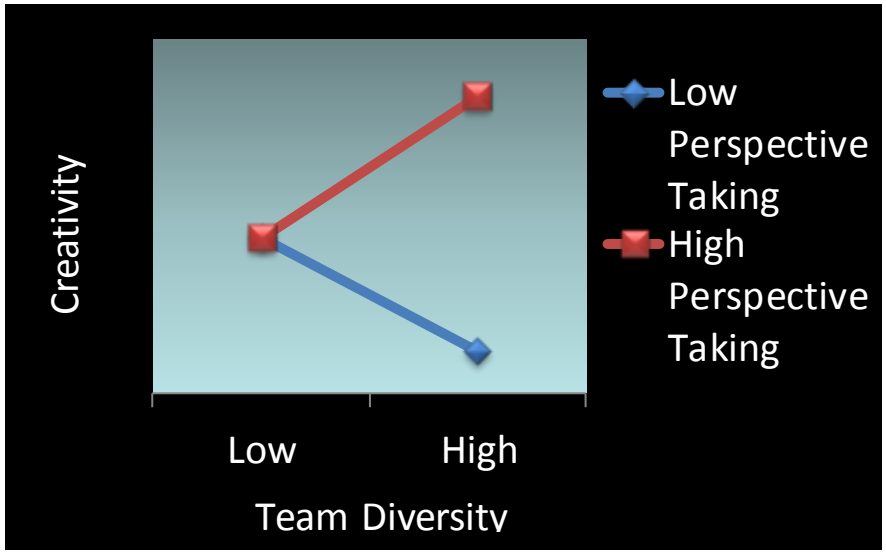
Hypothesis: The relationship between team diversity and team creativity is moderated by perspective taking (cf. Hoever et al., 2012, JAP).

Scenario 1: disordinal



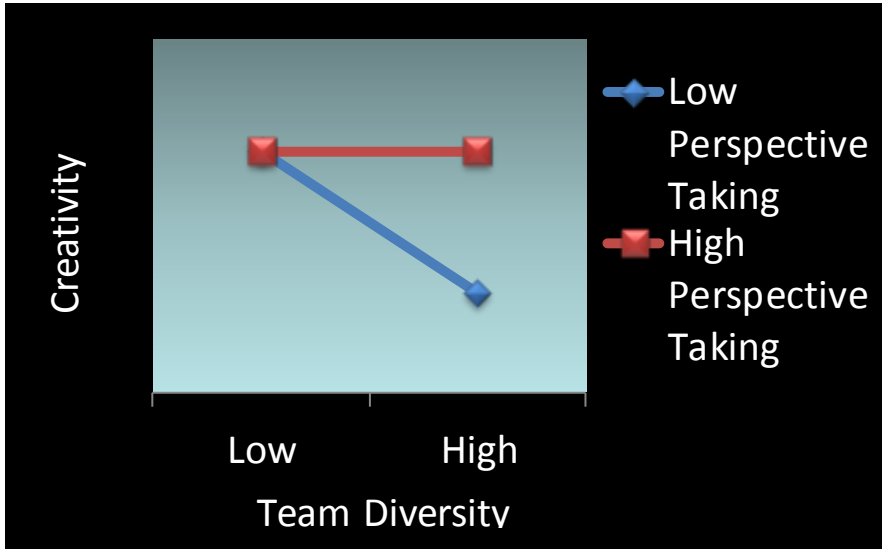
$Y = 0.00X + 0.00Z + 2.58*XZ + 2.54$

Scenario 2: ordinal

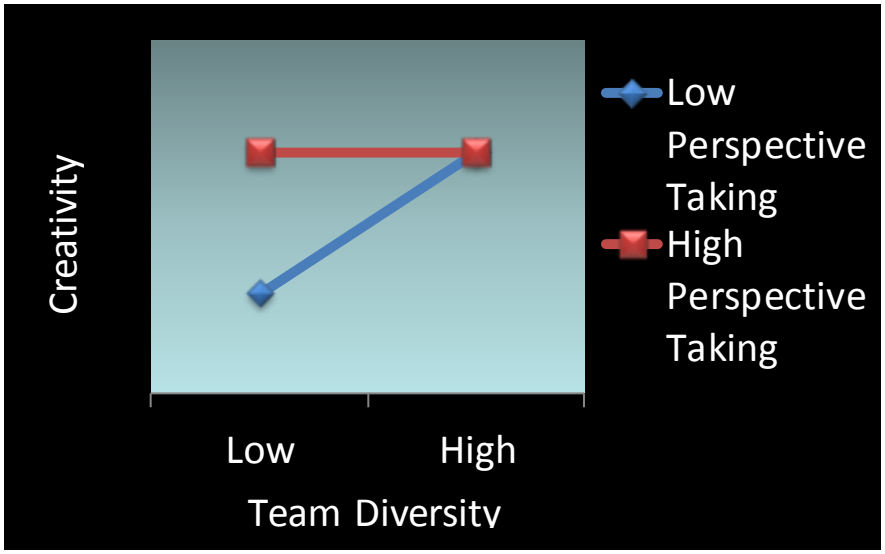


$Y = 0.00X + 1.50Z + 2.58*XZ + 2.54$

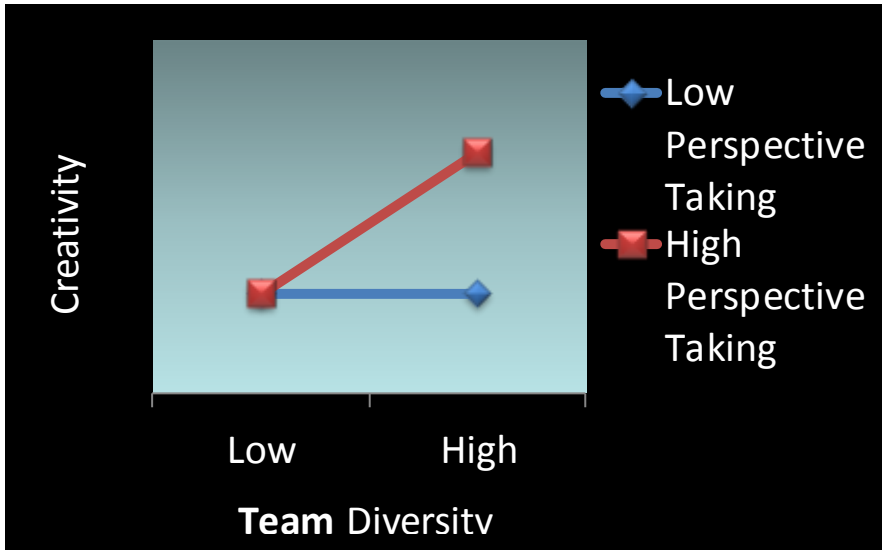
Scenario 1: buffering



Scenario 2: interference/antagonistic



Scenario 3: synergistic/enhancing



Testing two-way interactions in SPSS

- **Example data set of 424 employees**

- **Independent variables/moderators:**
  - Training, Autonomy, Responsibility, Age (all continuous)
- **Dependent variables:**
  - Job satisfaction, well being (continuous)
  - Receiving bonus (binary)
  - Days' absence in last year (count)

H1: Training has a more positive effect on job satisfaction for younger workers than for older workers

Testing two-way interactions in SPSS

- **IV: TRAIN_C**
- **Moderator: AGE_C**
- **DV: JOBSAT**

```
compute TRAXAGE = TRAIN_C*AGE_C.
```

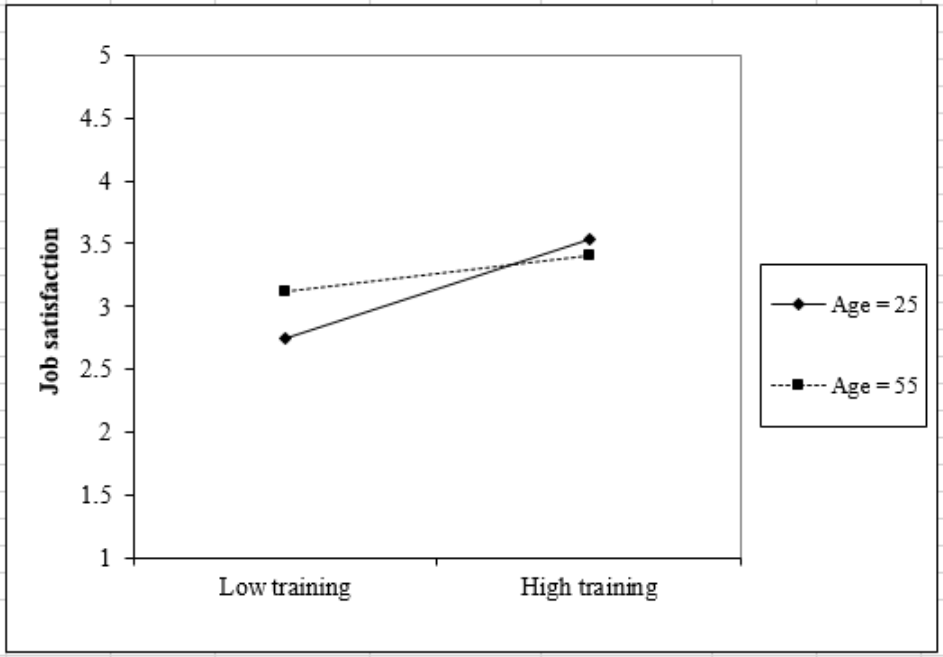
1. Compute interaction term

```
regression
 /statistics = r coeff bcov
 /dependent = JOBSAT
 /method = enter TRAIN_C AGE_C TRAXAGE.
```

2. Run regression to test moderation

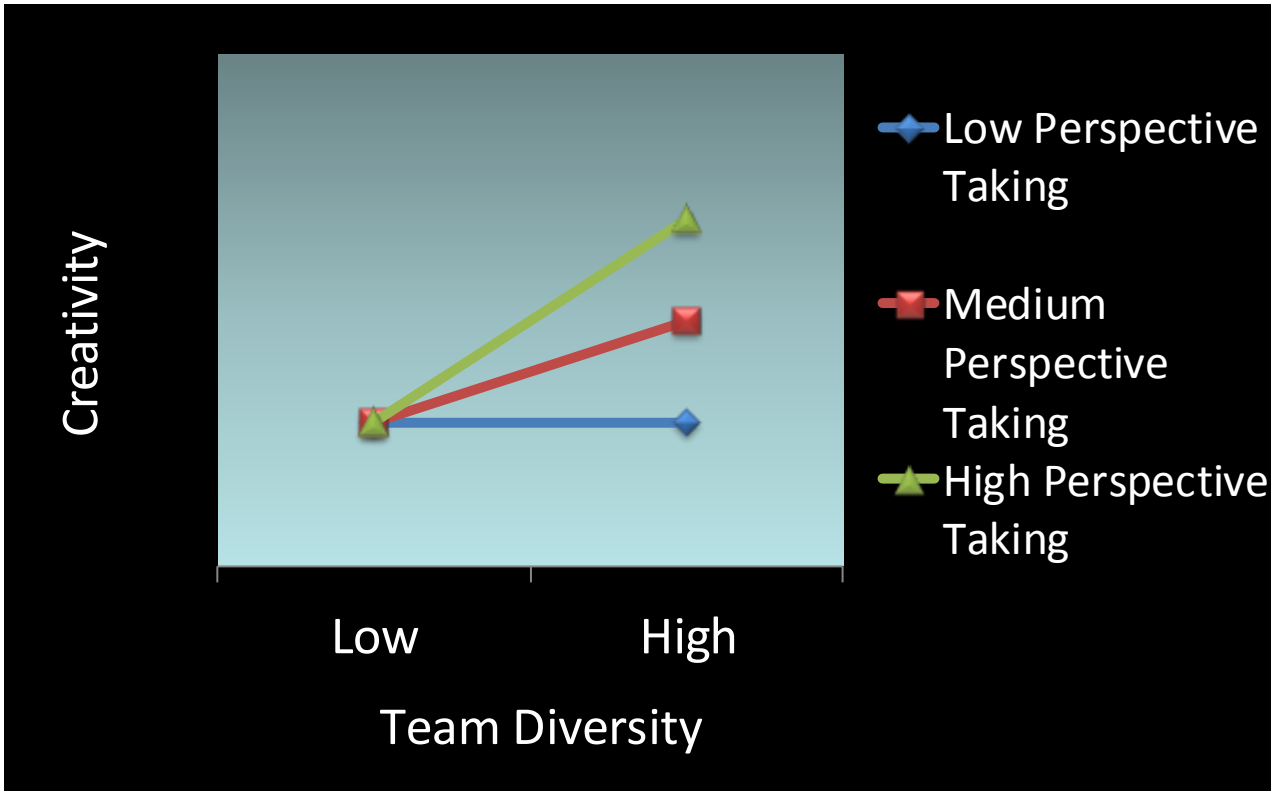
Plotting two-way interactions

2		
3	**Enter information from your regression in the shaded cells**	
4		
5	**Variable names:**	
6	Name of independent variable:	training
7	Name of moderator:	age
8		
9	**Unstandardised Regression Coefficients:**	
10	Independent variable:	0.38234
11	Moderator:	0.00413
12	Interaction:	-0.01276
13		
14	Intercept / Constant:	3.207
15		
16	**Means / SDs of variables:**	
17	Mean of independent variable:	0
18	SD of independent variable:	0.66045
19	Mean of moderator:	0
20	SD of moderator:	13.46
21		
22	**Values of variables at which to plot slopes*:**	
23	Low value of IV:	
24	High value of IV:	
25	Low value of moderator:	-16.55
26	High value of moderator:	13.45
27	*(* If left blank, this will automatically be done at one standard deviation above and below mean)*	
28		
29	**Optional alternative legend**:**	
30	Low value of independent variable:	
31	High value of independent variable:	
32	Low value of moderator:	Age = 25
33	High value of moderator:	Age = 55
34	*(** Leave blank for normal legend)*	



http://www.jeremydawson.co.uk/slopes.htm - "2-way with options" template

Probing two-way interactions: Simple slope tests (Aiken & West, 1991)



Simple slope tests: Direct method

5				
6	SIMPLE SLOPES ANALYSIS		Gradient of slope for Age = 25	0.593
7	Variance of coefficient of IV:	0.00269	t-value of slope for Age = 25	7.281
8	Variance of coefficient of interaction:	1.6E-05	p-value of slope for Age = 25	0.000
9	Covariance of coefficients of IV and interaction:	1.3E-05		
0			Gradient of slope for Age = 55	0.211
1	Sample size:	424	t-value of slope for Age = 55	2.736
2	Number of control variables:	0	p-value of slope for Age = 55	0.006
3				

These figures should be taken from the coefficient covariance matrix (acquired using the BCOV keyword in SPSS).

Note that the variance of a coefficient is the covariance of that coefficient with itself!

These are then produced automatically: here they tell us that the slope is positive and statistically significant at both 25 and 55 (although less at 55)

See Aiken & West (1991) or Dawson (2014) for formula

Simple slope tests: Indirect method

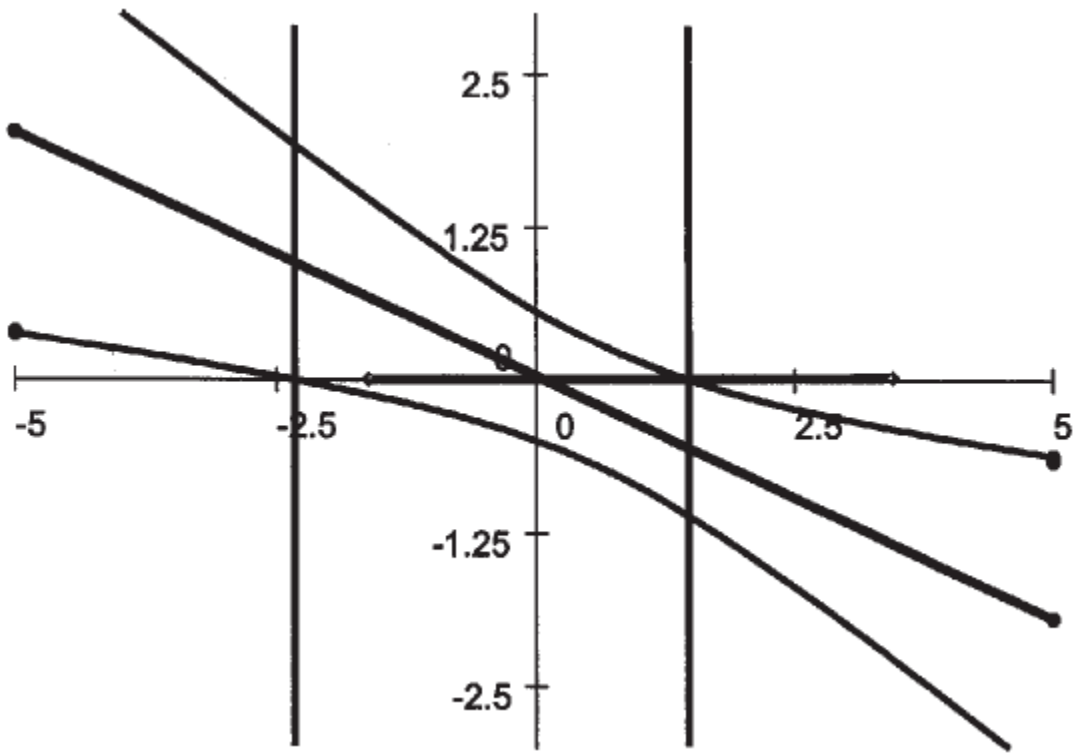
- **Principle:** The coefficient of the IV gives the slope when the moderator = 0
- **Method:** "Center" the moderator around the testing value; re-calculate interactions and run the regression
- **Interpretation:** The coefficient and p-value of the IV in the new analysis give the result of the simple slope test

```
compute AGE_55 = AGE-55.
compute TRAXAGE_55 = TRAIN_C*AGE_55.

regression
 /statistics=r coeff bcov
 /dependent=JOBSAT
 /method=enter TRAIN_C AGE_55 TRAXAGE_55.
```

Simple slope tests: Some thoughts

- Simple slope tests are far more meaningful when **meaningful** values of the moderator are used
- Ensure correct values are chosen **after centering** decision is made!
  - Here, for example, AGE was centered around the mean (41.55), so ages of 25 and 55 are actually -16.55 and 13.45 respectively
- Choosing values 1 SD above and below the mean is arbitrary and **should generally be avoided**
- Remember, statistical significance merely indicates a difference from zero – **it says nothing about the size or importance of an effect**

J-N regions of significance and confidence bands (Bauer & Curran, 2006)



Testing three-way interactions

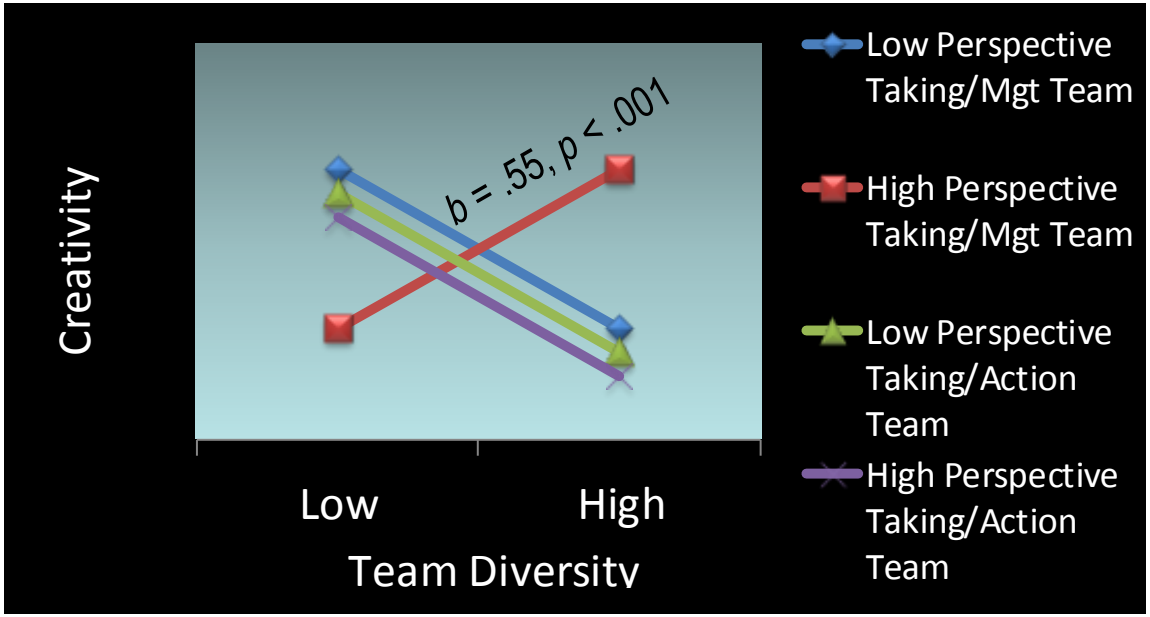
- $\hat{Y} = b_0 + b_1X + b_2Z + b_3W + b_4XZ + b_5XW + b_6ZW + b_7XZW$

Lower order effects

3-way interaction term

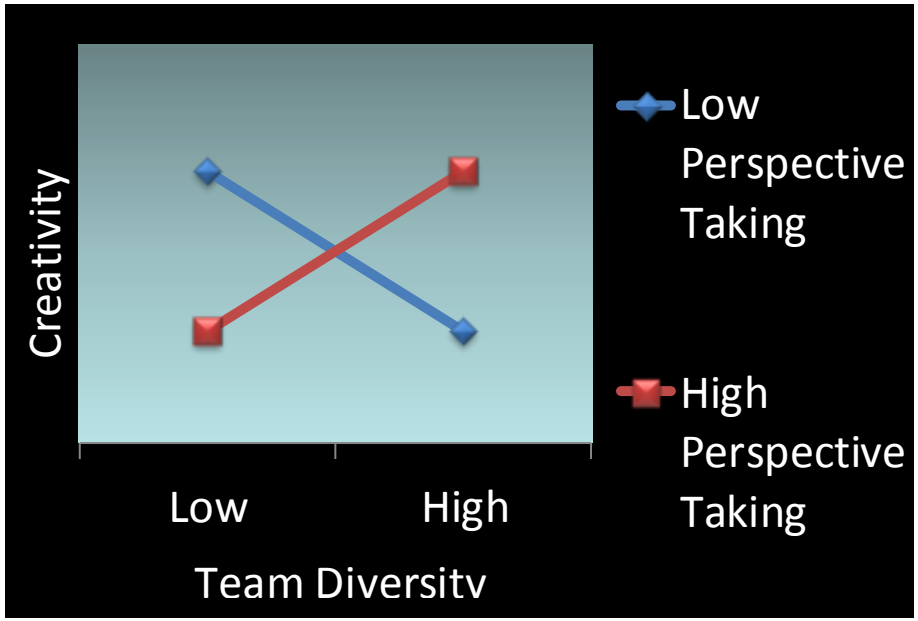
Probing three-way interactions: Simple slope tests (Aiken & West, 1991)

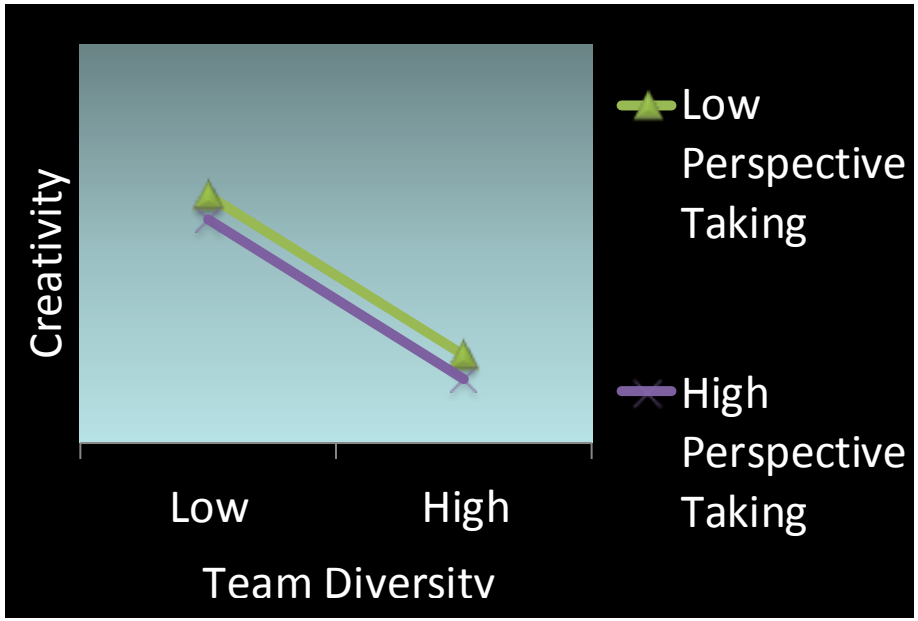
Hypothesis: The relationship between team diversity and team creativity is moderated by perspective taking for managerial teams.



Probing three-way interactions: Simple interaction tests (Aiken & West, 2000)

Hypothesis: The relationship between team diversity and team creativity is moderated by perspective taking for managerial, but not for action teams.

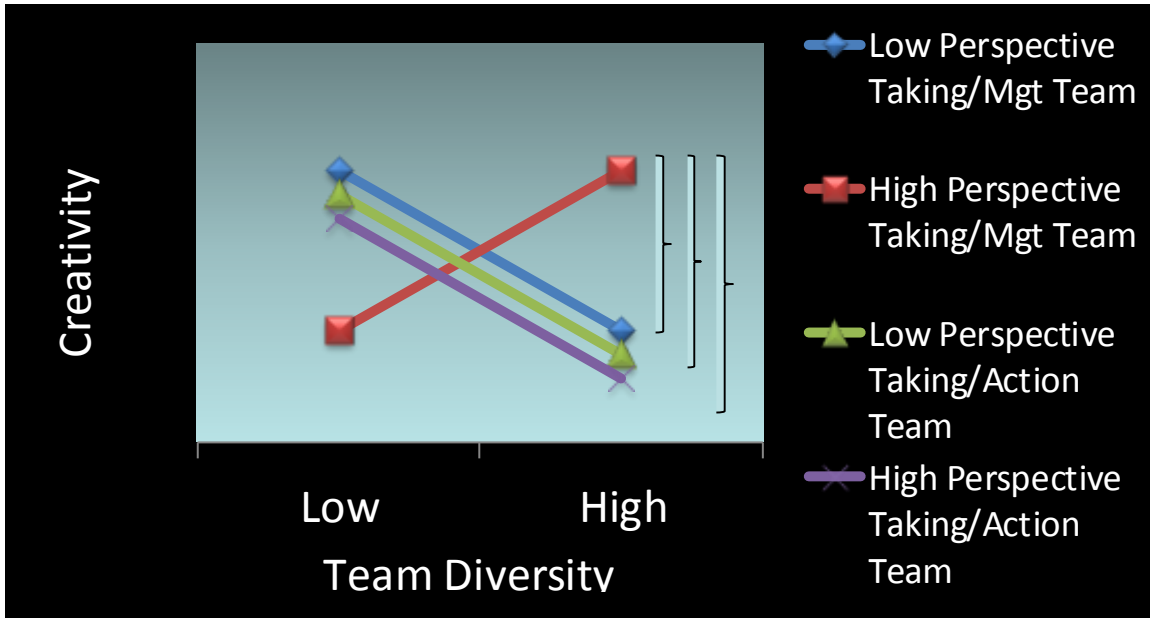




Managerial Teams **Action Teams**

Probing three-way interactions: Slope difference tests (Dawson & Richter, 2006)

Hypothesis: Team diversity predicts team creativity most strongly if teams use perspective taking and are managerial rather than action teams.



Testing three-way interactions

H2: The positive effect of training on job satisfaction for younger workers is strengthened when autonomy is higher

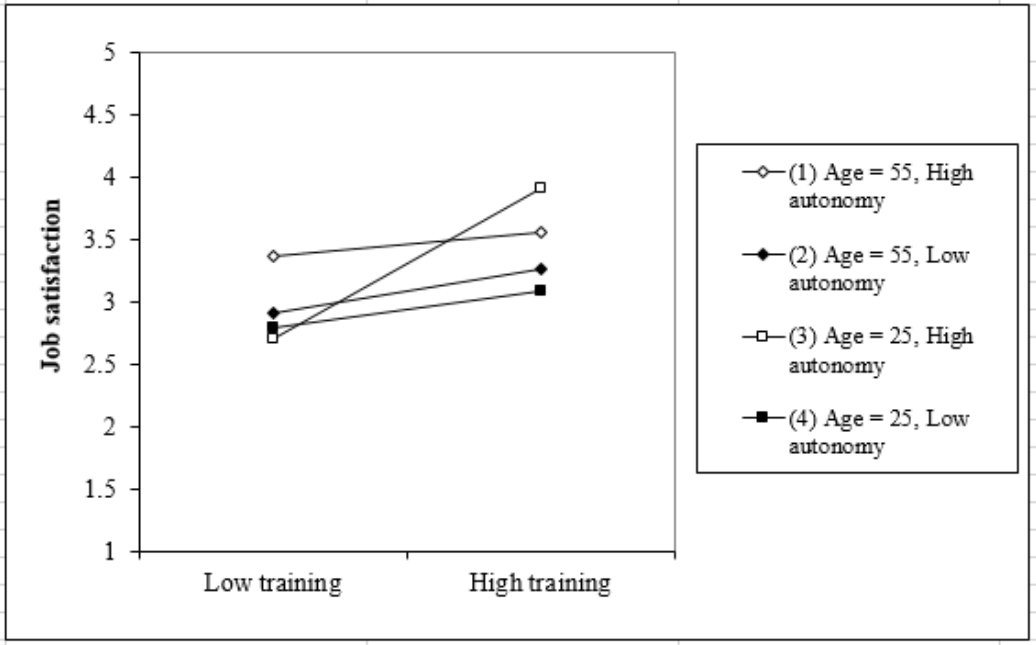
```
compute TRAXAUT = TRAIN_C*AUTON_C.
compute AUTXAGE = AUTON_C*AGE_C.
compute TRXAUXAG = TRAIN_C*AUTON_C*AGE_C.

regression
 /statistics=r coeff bcov
 /dependent=JOBSAT
 /method=enter TRAIN_C AUTON_C AGE_C
   TRAXAUT TRAXAGE AUTXAGE TRXAUXAG.
```

1. Compute remaining interaction terms

2. Run regression to test moderation

Plotting three-way interactions

4	**Enter information from your regression in the shaded cells**	
5		
6	**Variable names:**	
7	Name of variable 1:	training
8	Name of variable 2:	age
9	Name of variable 3:	autonomy
10		
11	**Unstandardised regression coefficients:**	
12	Var 1:	0.367
13	Var 2:	0.005
14	Var 3:	0.216
15	Var 1*Var 2:	-0.012
16	Var 1*Var 3:	0.141
17	Var 2*Var 3:	0
18	Var 1*Var 2*Var 3:	-0.016
19		
20	Intercept / Constant:	3.206
21		
22	**Means / SDs of variables:**	
23	Mean of Var 1:	0
24	Standard deviation of Var 1:	0.66045
25	Mean of Var 2:	0
26	Standard deviation of Var 2:	13.462
27	Mean of Var 3:	0
28	Standard deviation of Var 3:	0.85897
29		
30	**Values of variables at which to plot slopes*:**	
31	Var 1 - Low:	
32	Var 1 - High:	
33	Var 2 - Low:	-16.55
34	Var 2 - High:	13.45
35	Var 3 - Low:	
36	Var 3 - High:	
37	*(* If left blank, this will automatically be done at one standard deviation above and below mean)*	



Optional alternative legend:**	
Low value of Var 1:	
High value of Var 1:	
Low value of Var 2:	Age = 25
High value of Var 2:	Age = 55
Low value of Var 3:	
High value of Var 3:	
*(** Leave blank for normal legend)*	

http://www.jeremydawson.co.uk/slopes.htm - "3-way with options" template

Slope difference test

38		
39	**Additional information for slope difference test:**	
40	Sample size:	424
41	Number of control variables:	0
42		
43	Variance of Var1*Var2 coefficient:	0.000014
44	Variance of Var1*Var3 coefficent:	0.003486
45	Variance of Var1*Var2*Var3 coefficient:	0.000022
46	Covariance of Var1*Var2, Var1*Var3 coefficients:	0.000003
47	Covariance of Var1*Var2, Var1*Var2*Var3 coefficients:	-6.51E-07
48	Covariance of Var1*Var3, Var1*Var2*Var3 coefficients:	0.000022
49		

Slope difference tests:

Pair of slopes	t-value for slope difference	p-value for slope difference
(1) and (2)	-0.828	0.408
(1) and (3)	-4.771	0.000
(1) and (4)	-0.562	0.575
(2) and (3)	-3.982	0.000
(2) and (4)	0.311	0.756
(3) and (4)	4.296	0.000

These figures should be taken from the coefficient covariance matrix (acquired using the BCOV keyword in SPSS)

Be careful about the order: SPSS sometimes switches this around!

These are then produced automatically: here we find that slope 3 (age 25, high autonomy) is significantly greater than the other three slopes

It is important to hypothesize **which** slopes should be different from each other!

See Dawson & Richter (2006) or Dawson (2014) for formulas

End of section 1: Questions?

Session organizer

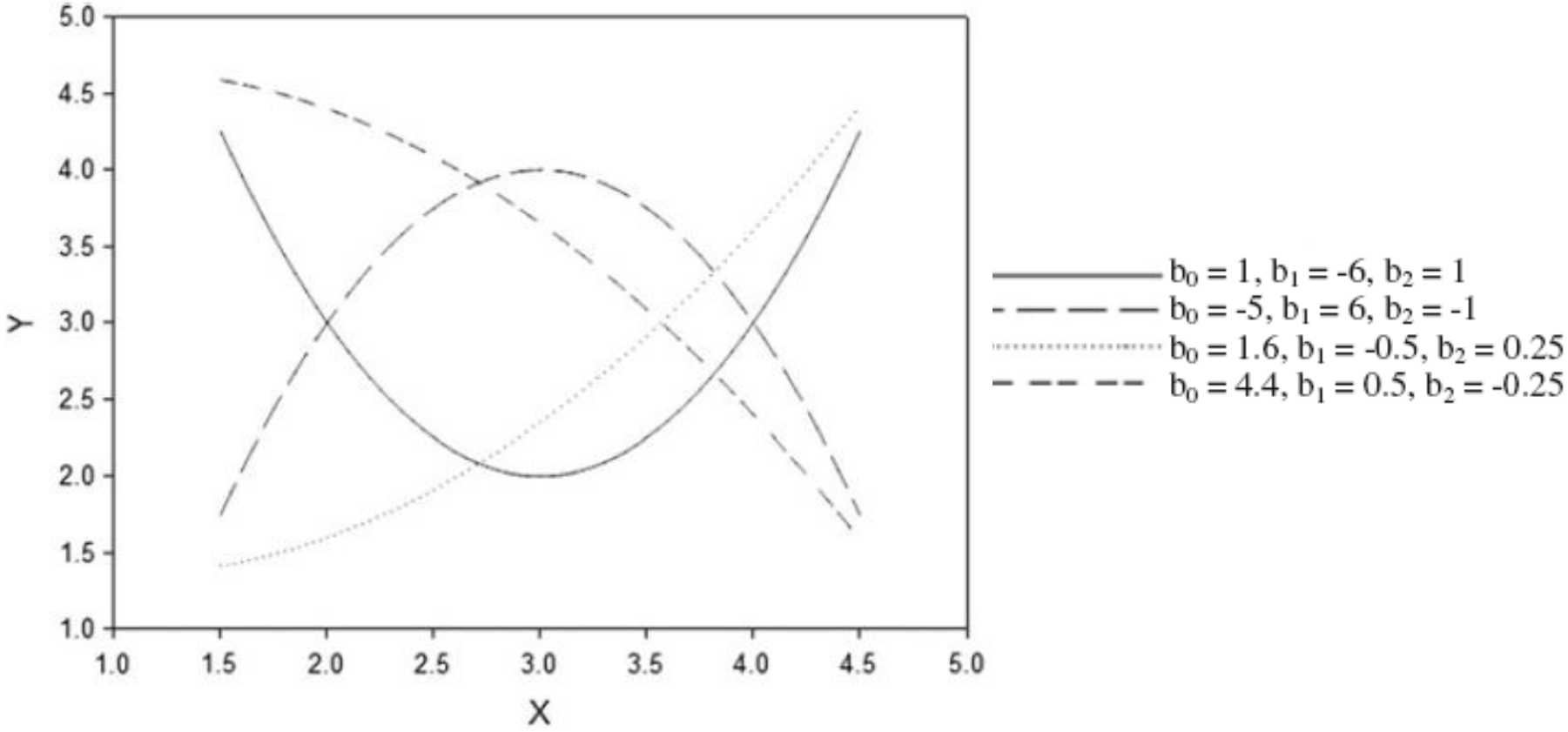
1. Testing and probing two-way and three-way interactions using MRA
2. <u>Curvilinear interactions</u>
3. Interactions with non-Normal outcomes
4. Extensions of MRA

Curvilinear interactions

- **Examples:**
  - **Baer & Oldham (2006, JAP): The curvilinear relationship between employees' experienced creative time pressure and creativity is moderated by amount of support for creativity.**
  - **Zhou et al. (2009, JAP): The curvilinear relationship between number of weak ties and creativity is moderated by conformity value.**

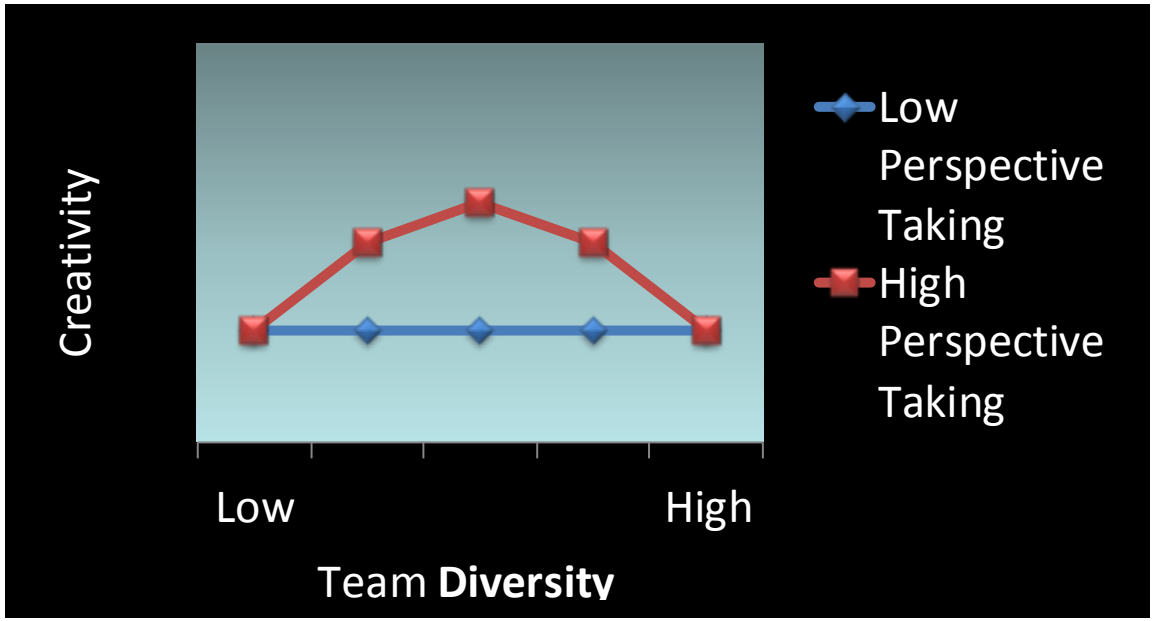
Curvilinear effects

- $\hat{Y} = b_0 + b_1X + b_2X^2$



Testing curvilinear interactions

Hypothesis: a curvilinear relationship between team diversity and team creativity moderated by perspective taking (cf. Hoever et al., 2012, JAP).



- $\hat{Y} = b_0 + b_1X + b_2X^2 + b_3Z + b_4XZ + + b_5X^2Z + r$

Testing a curvilinear relationship

H3: The relationship between responsibility and well-being is an inverted U shape: well-being is highest when responsibility is moderate

```
compute RESP_C2 = RESP_C*RESP_C.
```

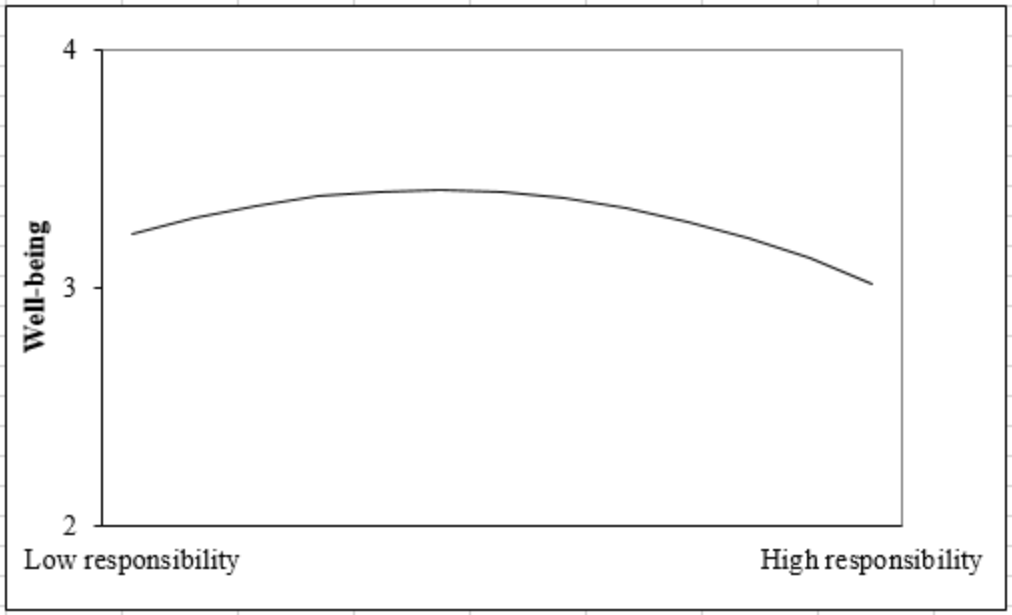
1. Compute quadratic (squared) term

```
regression
 /statistics=r coeff bcov
 /dependent=WELLBEING
 /method=enter RESP_C RESP_C2.
```

2. Run regression to test effect

Plotting a curvilinear relationship

Enter information from your regression in the shaded cells	
Variable names:	
Name of independent variable:	responsibility
Unstandardised Regression Coefficients:	
Independent variable:	-0.085
Independent variable squared:	-0.18
Intercept / Constant:	3.402
Means / SDs of IV:	
Mean:	0
SD:	0.82408



http://www.jeremydawson.co.uk/slopes.htm - "Quadratic regression" template

Testing a curvilinear interaction

H4: The relationship between responsibility and well-being is stronger when training is low

```
compute RESXTRA = RESP_C*TRAIN_C.
compute RES2XTRA = RESP_C2*TRAIN_C.

regression
/statistics=r coeff bcov
/dependent=WELLBEING
/method=enter RESP_C RESP_C2 TRAIN_C
 RESXTRA RES2XTRA.
```

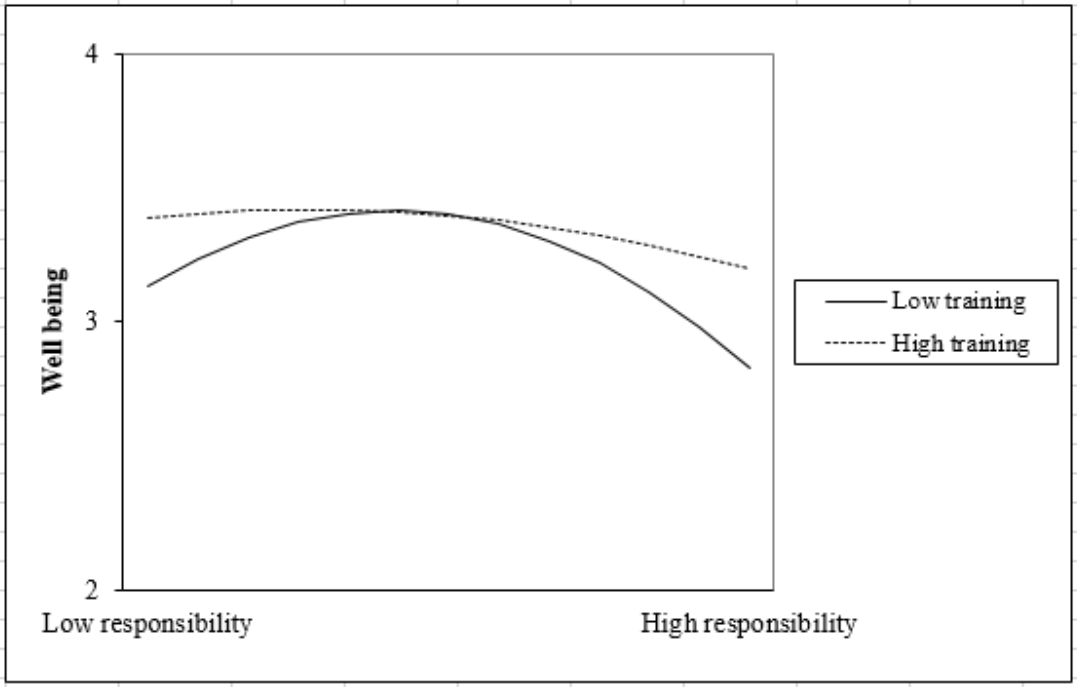
1. Compute **two** interaction terms

2. Run regression to test interaction

Note: Evidence of curvilinear interaction if and only if `RES2XTRA` *coefficient is significant*

Plotting a curvilinear interaction

3	**Enter information from your regression in the shaded cells**	
4		
5	**Variable names:**	
6	Name of independent variable:	responsibility
7	Name of moderator:	training
8		
9	**Unstandardised Regression Coefficients:**	
10	Independent variable:	-0.1
11	Independent variable squared:	-0.172
12	Moderator:	-0.002
13	Interaction - IV x Moderator:	0.034
14	Interaction - IV squared x Moderator:	0.156
15		
16	Intercept / Constant:	3.4
17		
18	**Means / SDs of variables:**	
19	Mean of independent variable:	0
20	SD of independent variable:	0.82408
21	Mean of moderator:	0
22	SD of moderator:	0.66045
23		
24		
25		
26		
27		



http://www.jeremydawson.co.uk/slopes.htm - "Quadratic two-way interactions"

Probing curvilinear interactions

- **Simple "slope" (or curve) test analogous to linear interactions, but with two versions:**
  i. Testing whether there is a *curvilinear* effect at a particular value of the moderator
  ii. Testing whether there is *any* effect at a particular value of the moderator

Probing curvilinear interactions (i)

Testing whether there is a *curvilinear* effect at a particular value of the moderator:

- Use indirect method of simple slope test and check IV^2 term
- e.g. for `TRAIN` = 4:

```
compute TRAIN_4=TRAIN-4.
compute RESXTRA_4 = RESP_C*TRAIN_4.
compute RES2XTRA_4 = RESP_C2*TRAIN_4.

regression
 /statistics=r coeff bcov
 /dependent=WELLBEING
 /method=enter RESP_C RESP_C2
  TRAIN_4 RESXTRA_4 RES2XTRA_4.
```

Check value/significance of this term

Probing curvilinear interactions (ii)

Testing whether there is *any* effect at a particular value of the moderator:

- Use indirect method of simple slope test and check for variance explained jointly by IV and IV^2 terms
- e.g. (having computed terms as on previous slide):

```
regression
 /statistics=r coeff bcov change
 /dependent=WELLBEING
 /method=enter TRAIN_4 RESXTRA_4 RES2XTRA_4
 /method = enter RESP_C RESP_C2 .
```

Need this keyword in syntax to give F-test

IV and IV^2 terms entered in separate (latter) step

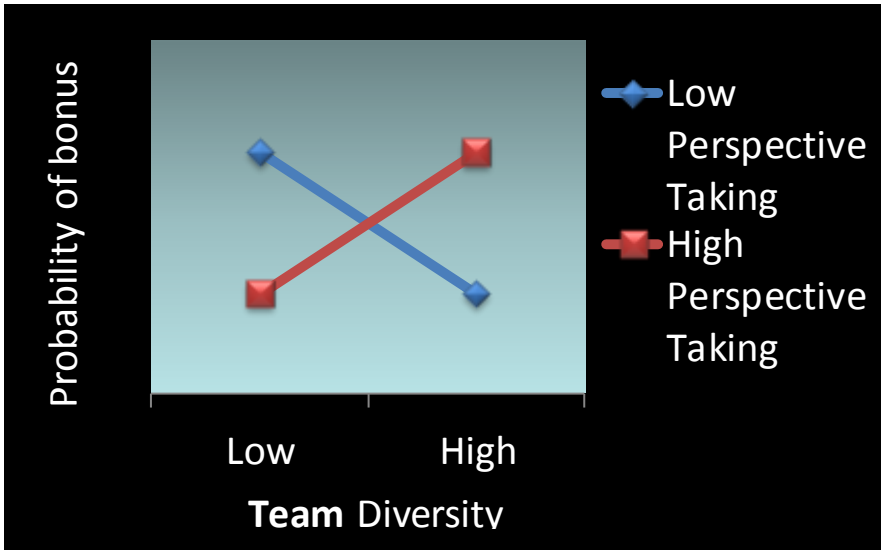
End of section 2: Questions?

Session organizer

1. Testing and probing two-way and three-way interactions using MRA
2. Curvilinear interactions
3. <u>Interactions with non-Normal outcomes</u>
4. Extensions of MRA

Interactions with Non-Normal outcomes

Hypothesis: The relationship between team diversity and receiving a team creativity bonus is moderated by perspective taking (cf. Hoever et al., 2012, JAP).



Testing interactions with binary outcomes

- ***Binary logistic* regression**

- **Logit $(\hat{Y}) = b_0 + b_1X + b_2Z + b_3XZ$**

Logit link function

Note: Logit($\hat{Y}$) = ln[$\hat{Y}$/(1- $\hat{Y}$)]

Testing an interaction with a binary outcome

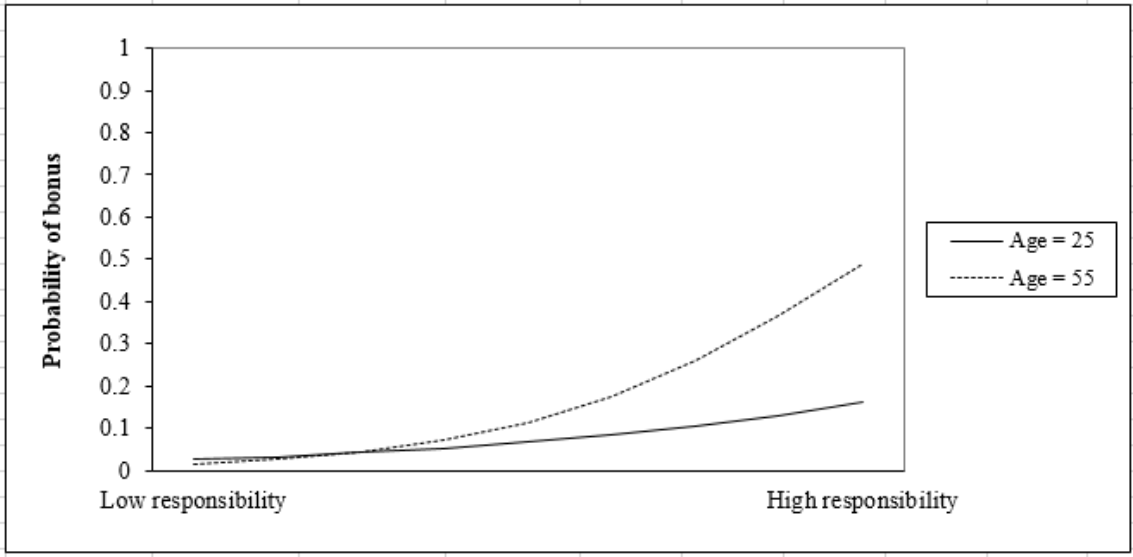
H5: Employees with more responsibility are more likely to receive a bonus when they are older

```
logistic regression variables BONUS
 /method = enter RESP_C AGE RESP_C*AGE.
```

Logistic regression syntax: no need to compute interaction term separately!

Plotting an interaction with a binary outcome

2		
3	**Enter information from your logistic regression in the shaded cells**	
4		
5	**Variable names:**	
6	Name of independent variable:	responsibility
7	Name of moderator:	age
8		
9	**Unstandardised Regression Coefficients:**	
10	Independent variable:	0.15
11	Moderator:	0.019
12	Interaction:	0.042
13		
14	Constant:	-3.115
15		
16	**Mean/SD of IV:**	
17	Mean of independent variable:	0
18	SD of independent variable:	0.82408
19		
20	**Values of moderator at which to plot slopes:**	
21	Low:	25
22	High:	55
23		
24		
25	**Optional alternative legend*:**	
26	Low value of independent variable:	
27	High value of independent variable:	
28	Low value of moderator:	Age = 25
29	High value of moderator:	Age = 55
30	*(* Leave blank for normal legend)*	
31		



http://www.jeremydawson.co.uk/slopes.htm - "2-way logistic interactions"

Probing interactions with non-normal outcomes

- **Simple "slope" tests need to be done using the indirect method**
- e.g. for AGE = 25:

```
compute AGE_25 = AGE-25.

logistic regression variables BONUS
 /method = enter RESP_C AGE_25 RESP_C*AGE_25.
```

Check value/significance of this term

Testing interactions with discrete (count) outcomes

- ***Poisson* or *Negative Binomial* regression**

- $\text{Log}\ (\hat{Y}) = b_0 + b_1 X + b_2 Z + b_3 XZ$

Natural log link function

Testing an interaction with a count outcome

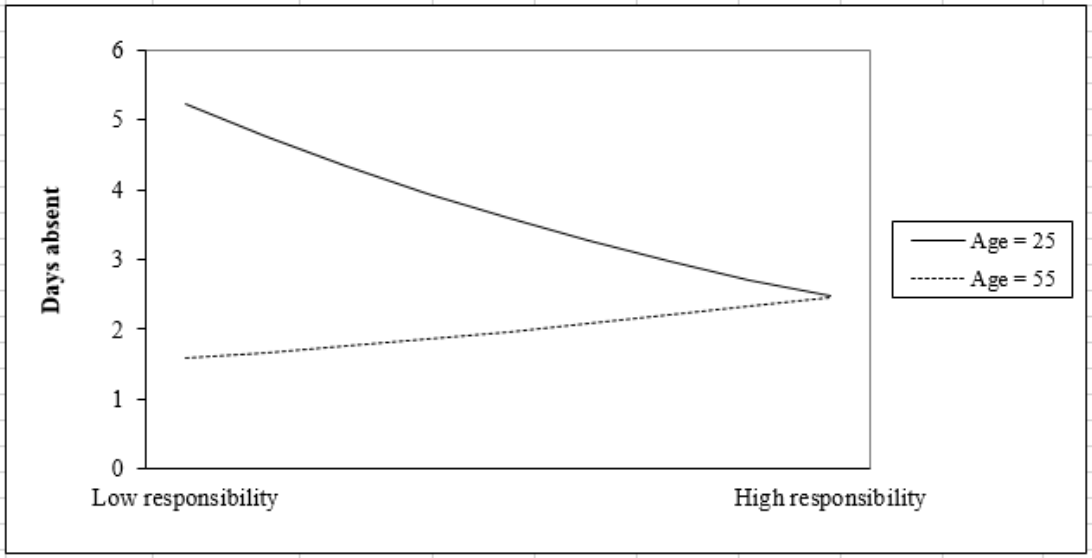
H6: Employees with less responsibility are likely to have more days' absence when they are younger

```
genlin ABSENCE with RESP_C AGE
 /model RESP_C AGE RESP_C*AGE
 intercept = yes distribution = poisson
 link = log.
```

Generalized linear models syntax: no need to compute interaction term separately!

Plotting an interaction with a count outcome

Enter information from your logistic regression in the shaded cells	
Variable names:	
Name of independent variable:	responsibility
Name of moderator:	age
Unstandardised Regression Coefficients:	
Independent variable:	-1.055
Moderator:	-0.02
Interaction:	0.024
Constant:	1.779
Mean/SD of IV:	
Mean of independent variable:	0
SD of independent variable:	0.82408
Values of moderator at which to plot slopes:	
Low:	25
High:	55
Optional alternative legend*:	
Low value of independent variable:	
High value of independent variable:	
Low value of moderator:	Age = 25
High value of moderator:	Age = 55
(Leave blank for normal legend)*	



http://www.jeremydawson.co.uk/slopes.htm - "2-way Poisson interactions"

End of section 3: Questions?

Session organizer

1. Testing and probing two-way and three-way interactions using MRA
2. Curvilinear interactions
3. Interactions with non-Normal outcomes
4. <u>Extensions of MRA</u>

Cross-level interactions

-level-1 model specification

Intercept and slope for each group *j*

Residual

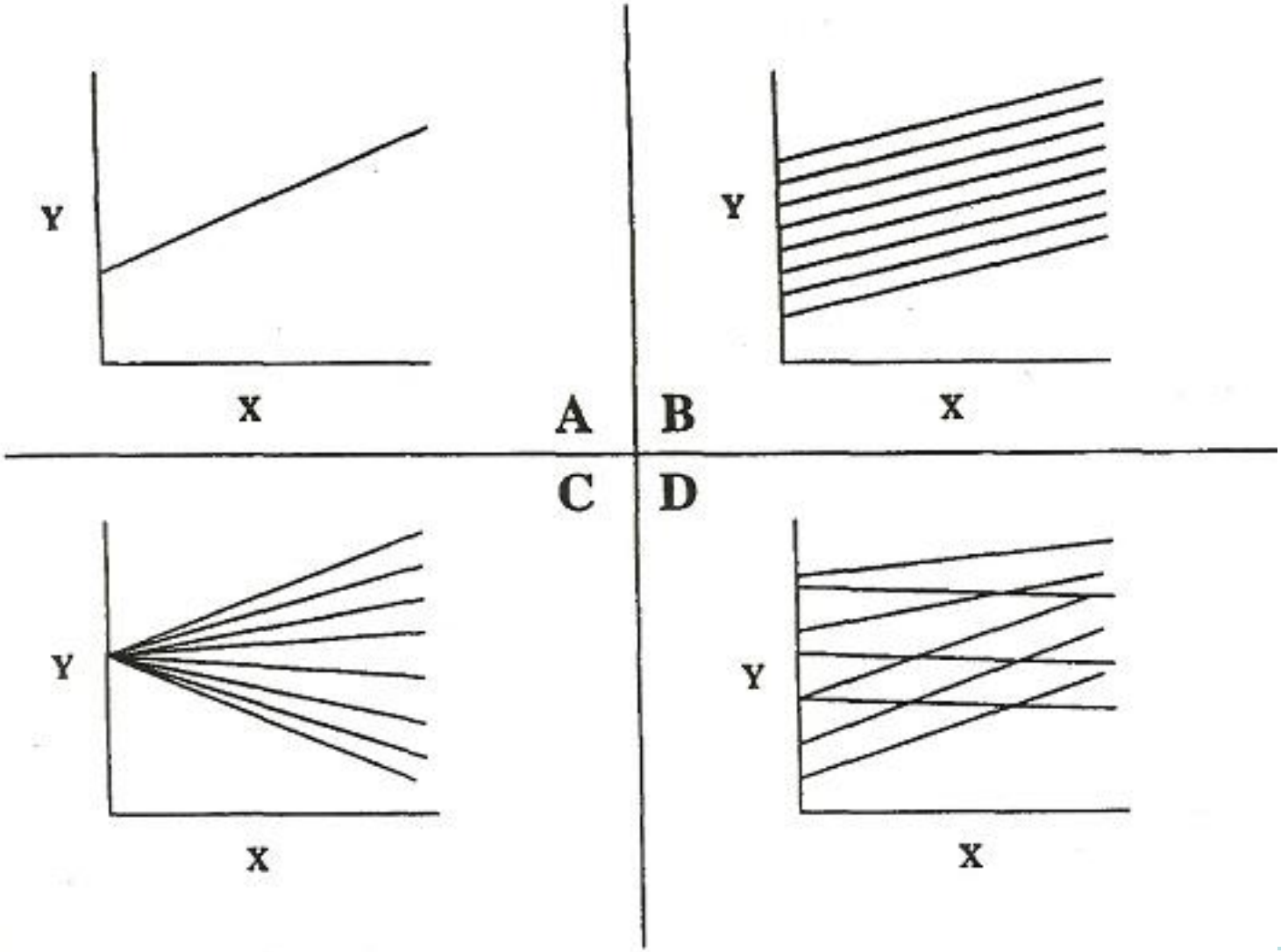
- Level- 1: $Yij = \beta_{0j} + \beta_{1j}X_{ij} + rij$

Predictor value of individual *i* in group *j*

Outcome measure for individual *i* in group *j*

Cross-level interactions

-level-2 model specification



Cross-level interactions

-level-2 model specification

- Level- 1: $Yij = \beta_{0j} + \beta_{1j}X_{ij} + rij$

Group level variable

- Level- 2: $\beta_{0j} = \gamma_{00} + \gamma_{01}G_j + U_{0j}$

Second stage intercept terms

Slopes relating *Gj* to intercept and slope terms from level 1 equation

Level 2 residuals

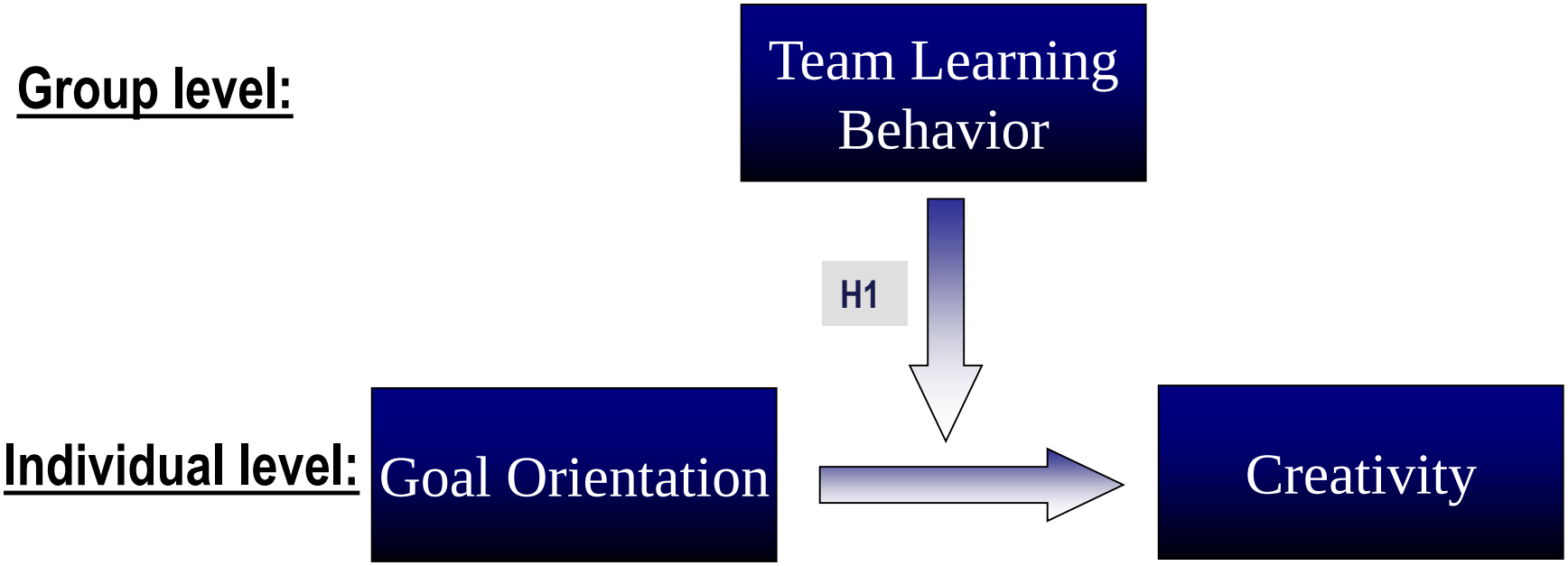
$$\beta_{1j} = \gamma_{10} + \gamma_{11}G_j + U_{1j}$$

Multilevel analysis

-hypotheses

Hirst et al. (2008, AMJ):

H1: Team learning behavior (GL) moderates the goal orientation (IL) — creativity (IL) relationship



Probing multilevel interactions

- Interactions can be plotted using the **same template** as relevant for single-level interactions
  - Estimates produced in output are equivalent to unstandardized coefficients in ordinary regression
  - Care is needed over mean & SD of variables
- However, in general, simple slope & slope difference tests **do not** work
- Simple slope tests can be done instead using the **indirect method**
- Slope difference tests are more complicated!

Interactions in SEM

- Mplus allows interactions between latent variables
  - All latent variables have mean & SD fixed at 0 and 1
  - Intercept given by weighted mean of intercepts of indicator variables for DV
- Simple slope tests cannot be conducted, however
  - Given the (relatively) arbitrary nature of the latent variables, it is doubtful whether they would be meaningful in any case!

Testing multiple interactions

- Best to do this simultaneously
- Difficult to plot, however
- If multiple two-way interactions, but involving no more than three variables, can do it via the 3-way template, leaving unused coefficients as 0

- **Always consider what is necessary to test your specific hypothesis!**

End of PDW: Questions?