Some Issues When Analyzing the Impact of Interaction

Dr Jian Liu

Brock University

Agenda

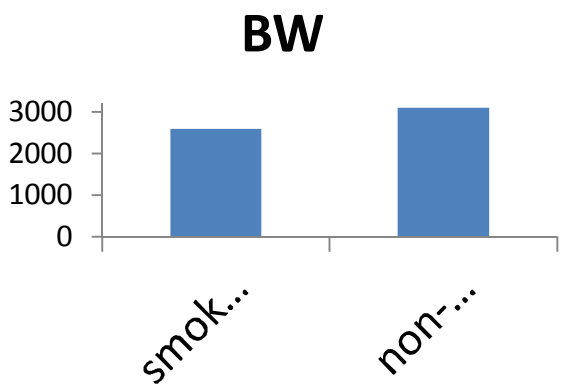
- Definition
- Examples
- Approaches

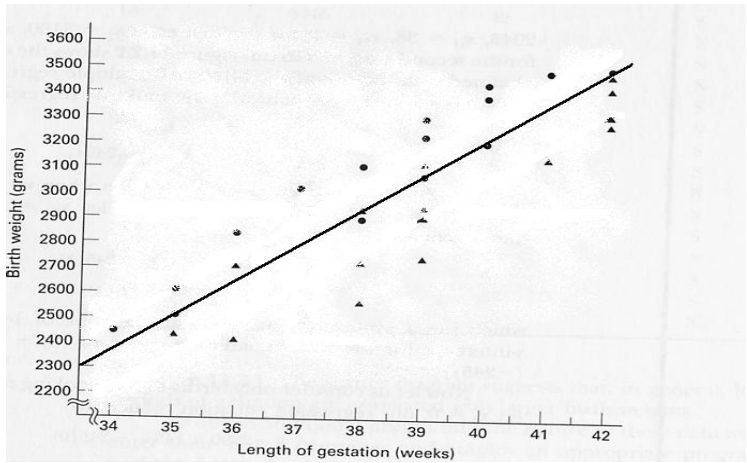
What is interaction?

Definition: there is an interaction between two factors if the ***effect*** of one of the factors changes for different categories of the other factor.

Examples

- Maternal smoking and gestational age on birth weight
  - Maternal smoking -> lower birth weight
  - Birth weight is positively associated with gestational age





No interaction!

Take both maternal smoking and gestational age into account

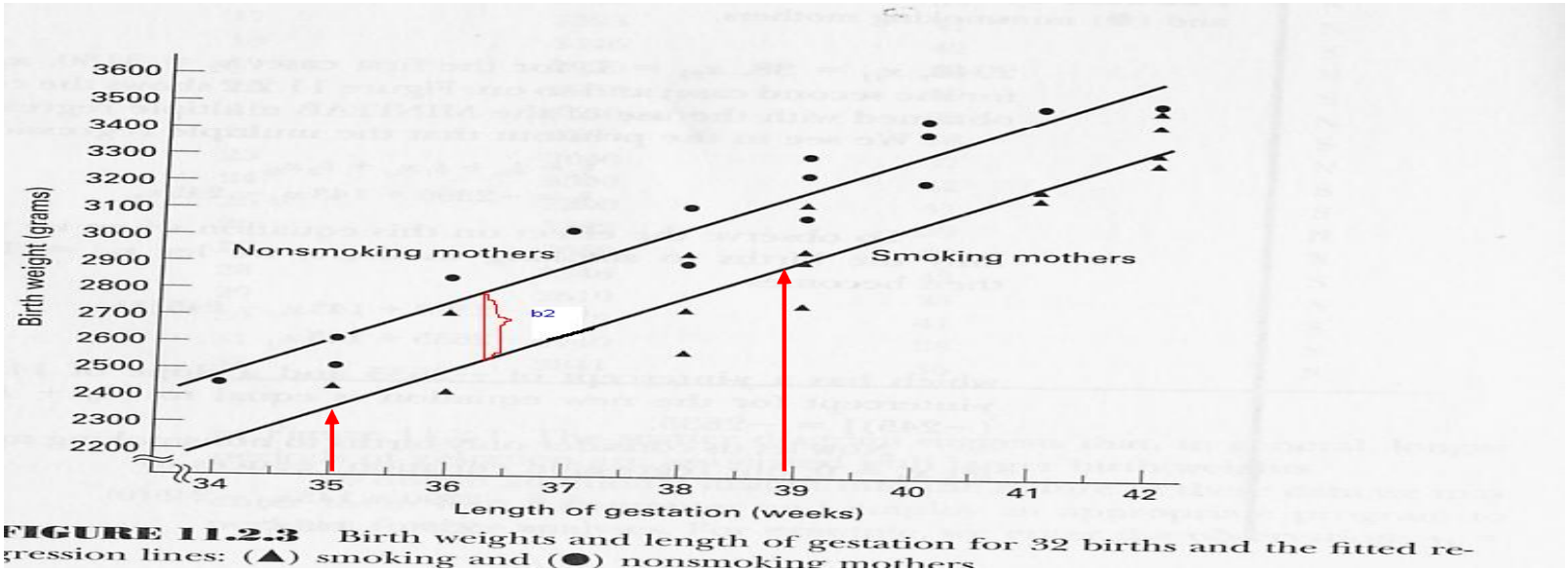
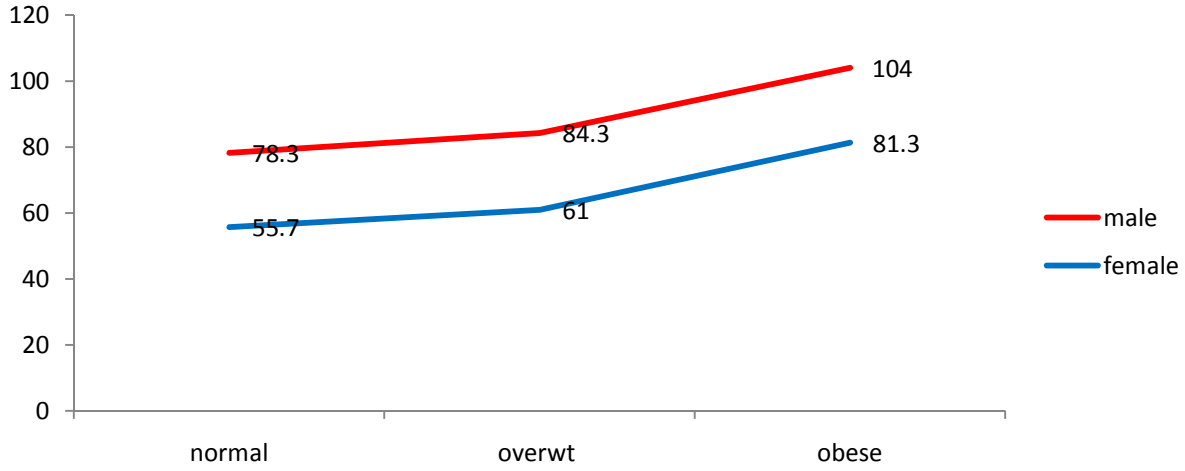


FIGURE 11.2.3 Birth weights and length of gestation for 32 births and the fitted regression lines: (▲) smoking and (●) nonsmoking mothers

The effect of maternal smoking on BW does not change by the gestational age

Waist circumference (cm) by obesity status and sex

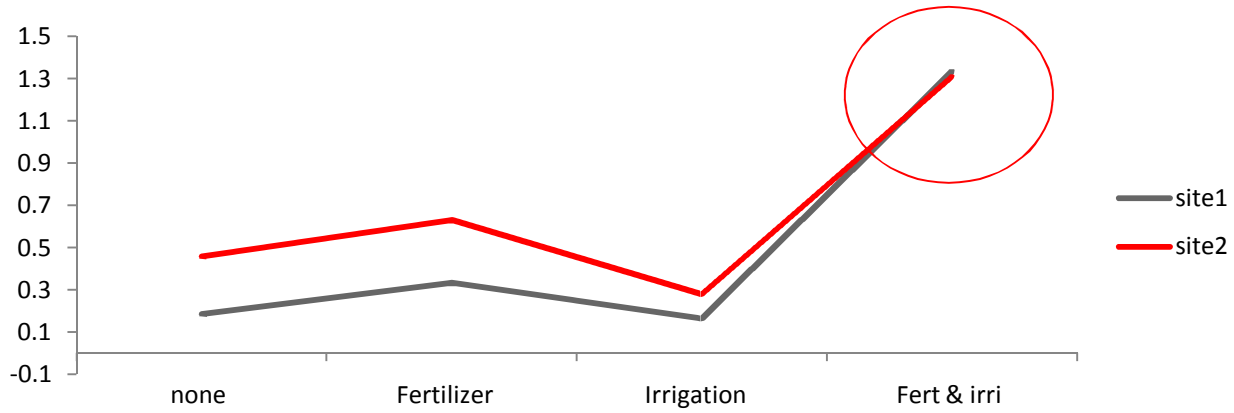
	Obesity status (defined by BMI)		
	normal	*overwt*	*obese*
male	*78.3*	*84.3*	*104.0*
female	*55.7*	*61.0*	*81.3*



No interaction!

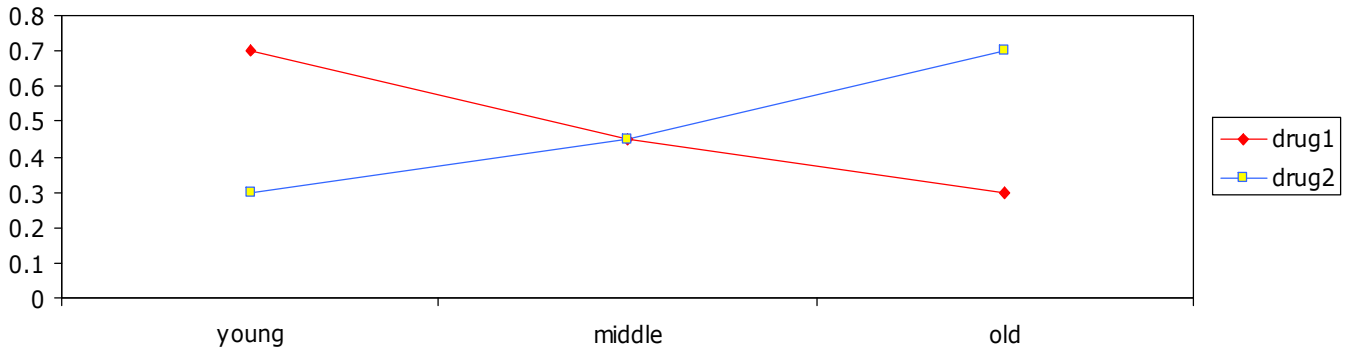
Interaction

	none	Fertilizer	Irrigation	Fert & irri
site1	.184	.332	.164	1.334
site2	.458	.630	.278	1.308



Interaction

	The percentage (%) of	Effective treatment
Age group	DrugA	DrugB
Young	0.70	0.30
Middle	0.45	0.45
Old	0.30	0.70



The treatment effect of drugs depends on age, so there is a interaction between drugs and age group

Approaches

- Two approaches can be used when there is an interaction
  - Comparison separately

	The percentage (%) of	Effective treatment
Age group	DrugA	DrugB
Young	0.70	0.30
Middle	0.45	0.45
Old	0.30	0.70

  - Multivariate analysis

Example: The change in serum cholesterol level (mg/dL) after 3 months trial

A clinical trial is designed to compare the effect of a standardized low-fat diet with a new drug lowering serum cholesterol in middle-age men. A placebo group receives an inert substance resembling the drug in all outward respects.

---------- treat=diet ----------

The MEANS Procedure

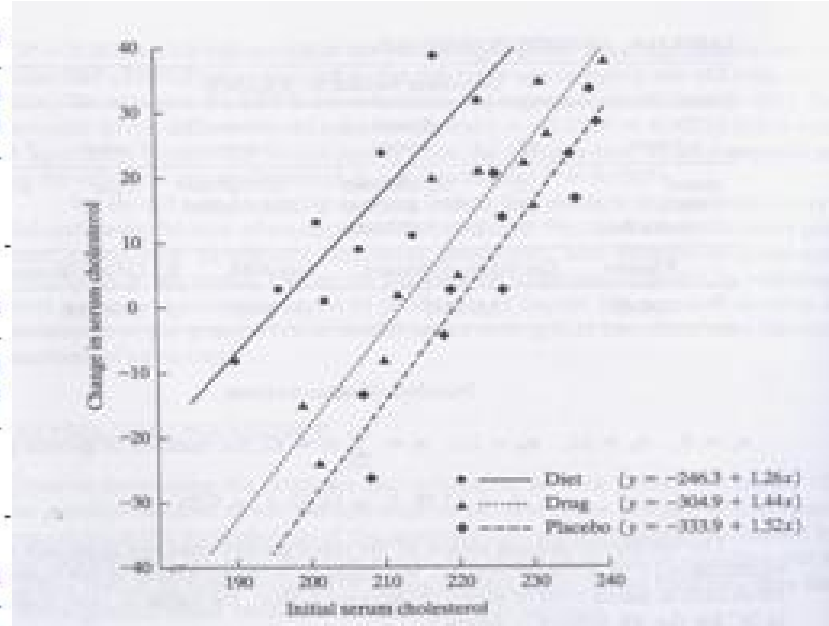
Variable	N	Mean	Std Dev	Minimum	Maximum
SC_1st	9	208.0000000	10.0000000	196.0000000	229.0000000
SC_chge	9	13.7777778	15.2379278	-8.0000000	39.0000000

---------- treat=drug ----------

Variable	N	Mean	Std Dev	Minimum	Maximum
SC_1st	12	219.9166667	13.3311551	199.0000000	240.0000000
SC_chge	12	11.3333333	19.9468992	-24.0000000	38.0000000

---------- treat=placebo ----------

Variable	N	Mean	Std Dev	Minimum	Maximum
SC_1st	11	225.1818182	11.3385906	207.0000000	239.0000000
SC_chge	11	9.2727273	18.5261486	-26.0000000	34.0000000



SC_chge

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	100.470960	50.235480	0.15	0.8608
Error	29	9666.404040	333.324277		
Corrected Total	31	9766.875000			

R-Square	Coeff Var	Root MSE	SC_chge Mean
0.010287	161.3894	18.25717	11.31250

Source	DF	Type I SS	Mean Square	F Value	Pr > F
gp	2	100.4709596	50.2354798	0.15	0.8608

Dependent Variable: SC_1st

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	1506.946970	753.473485	5.41	0.0101
Error	29	4040.553030	139.329415		
Corrected Total	31	5547.500000			

R-Square	Coeff Var	Root MSE	SC_1st Mean
0.271644	5.405284	11.80379	218.3750

Source	DF	Type I SS	Mean Square	F Value	Pr > F
gp	2	1506.946970	753.473485	5.41	0.0101

Adjusted for the 1st measurement of serum cholesterol

The GLM Procedure

Dependent Variable: SC_chge

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	5154.096971	1718.032324	10.43	<.0001
Error	28	4612.778029	164.742072		
Corrected Total	31	9766.875000			

R-Square	Coeff Var	Root MSE	SC_chge Mean
0.527712	113.4602	12.83519	11.31250

Source	DF	Type I SS	Mean Square	F Value	Pr > F
gp	2	100.470960	50.235480	0.30	0.7396
SC_1st	1	5053.626011	5053.626011	30.68	<.0001

Source	DF	Type III SS	Mean Square	F Value	Pr > F
gp	2	2070.083620	1035.041810	6.28	0.0056
SC_1st	1	5053.626011	5053.626011	30.68	<.0001

Adding interaction term

The GLM Procedure

Dependent Variable: SC_chge

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	7025.015858	1405.003172	13.32	<.0001
Error	26	2741.859142	105.456121		
Corrected Total	31	9766.875000			

R-Square	Coeff Var	Root MSE	SC_chge Mean
0.719270	90.77731	10.26918	11.31250

Source	DF	Type I SS	Mean Square	F Value	Pr > F
gp	2	100.470960	50.235480	0.48	0.6264
SC_1st	1	5053.626011	5053.626011	47.92	<.0001
SC_1st*gp	2	1870.918887	935.459443	8.87	0.0012

Source	DF	Type III SS	Mean Square	F Value	Pr > F
gp	2	2072.875499	1036.437750	9.83	0.0007
SC_1st	1	2843.384014	2843.384014	26.96	<.0001
SC_1st*gp	2	1870.918887	935.459443	8.87	0.0012

Notes

- Using R^2 (adjusted R^2) to evaluate if it is necessary

Adjusted R^2=
1- (1-R^2)(n-1)/(n-[k+1])

SC_chge

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	100.470960	50.235480	0.15	0.8608
Error	29	9666.404040	333.324277		
Corrected Total	31	9766.875000			

R-Square	Coeff Var	Root MSE	SC_chge Mean
0.010287	161.3894	18.25717	11.31250

Dependent Variable: SC_chge

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	5154.096971	1718.032324	10.43	<.0001
Error	28	4612.778029	164.742072		
Corrected Total	31	9766.875000			

R-Square	Coeff Var	Root MSE	SC_chge Mean
0.527712	113.4602	12.83519	11.31250

Dependent Variable: SC_chge

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	7025.015858	1405.003172	13.32	<.0001
Error	26	2741.859142	105.456121		
Corrected Total	31	9766.875000			

R-Square	Coeff Var	Root MSE	SC_chge Mean
0.719270	90.77731	10.26918	11.31250

Notes

- If interaction term is not significant (adjusted R^2 will decrease), the final model should not include the interaction term.

Thank you!