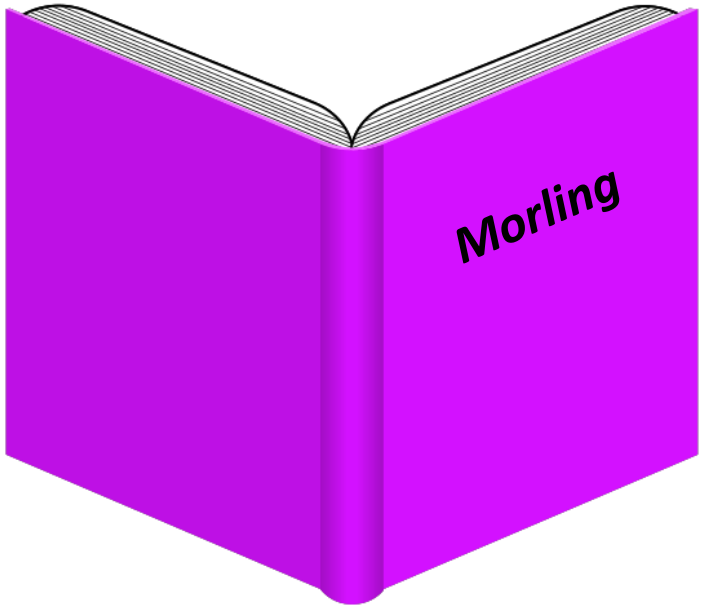




Research Methods and Statistics

Lecture 13: Factorial design

Riet van Bork

Pictures source: pixabay.org

Book chapters

- Chapter 4: Ethical guidelines for Psychology Research
- Chapter 12: Experiments with more than one independent variable
- Chapter 4 is not less important, but is explained well in the book. Remember that chapters not discussed in the lecture are still exam material.
- In this lecture we also make a start on some content of Chapter 11: obscuring factors



Milgram experiment 1963

How to balance risk to participants with benefit to society? In interviews years after the Milgram experiment, participants reported worrying for weeks about the learner's welfare.

Experiments

Strategies for relaxation/stress relief



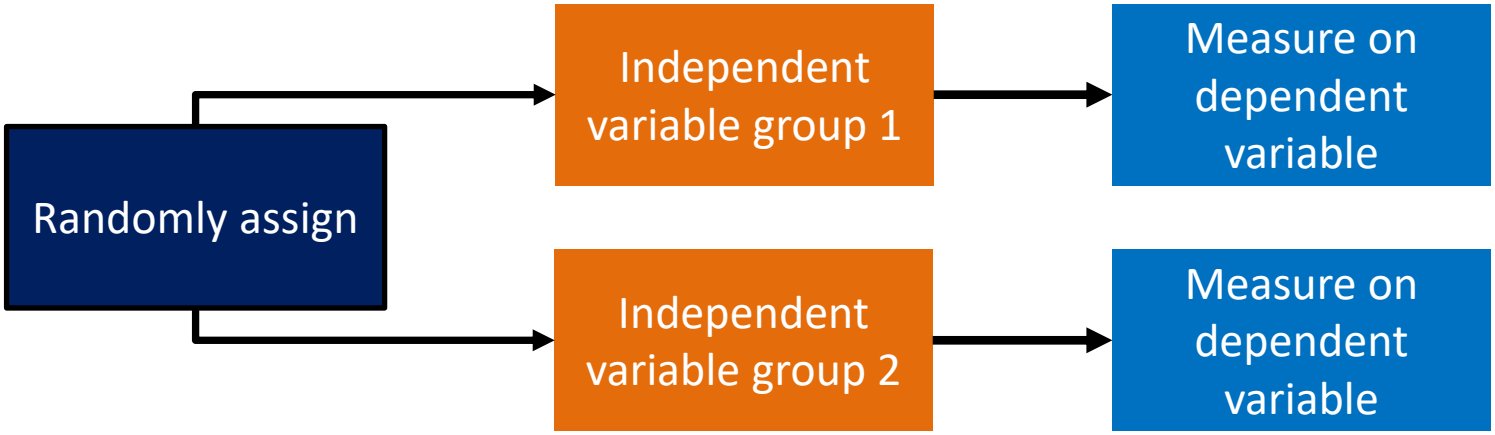


Picture source: pixabay.com

Experiments

How to study whether Thai boxing is an effective stress relief strategy?

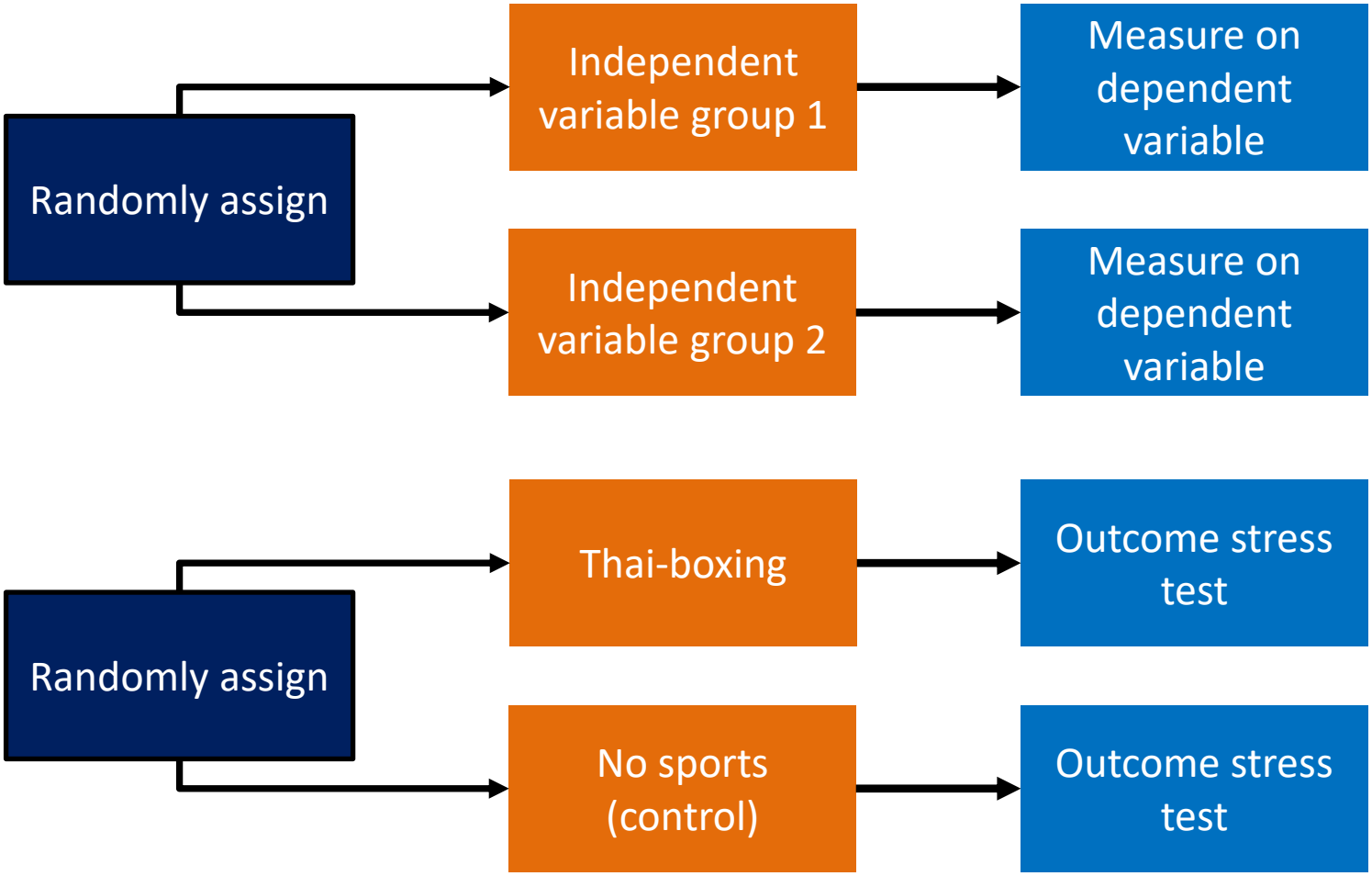
Chapter 10: independent-groups design (posttest only):



Experiments

How to study whether Thai boxing is an effective stress relief strategy?

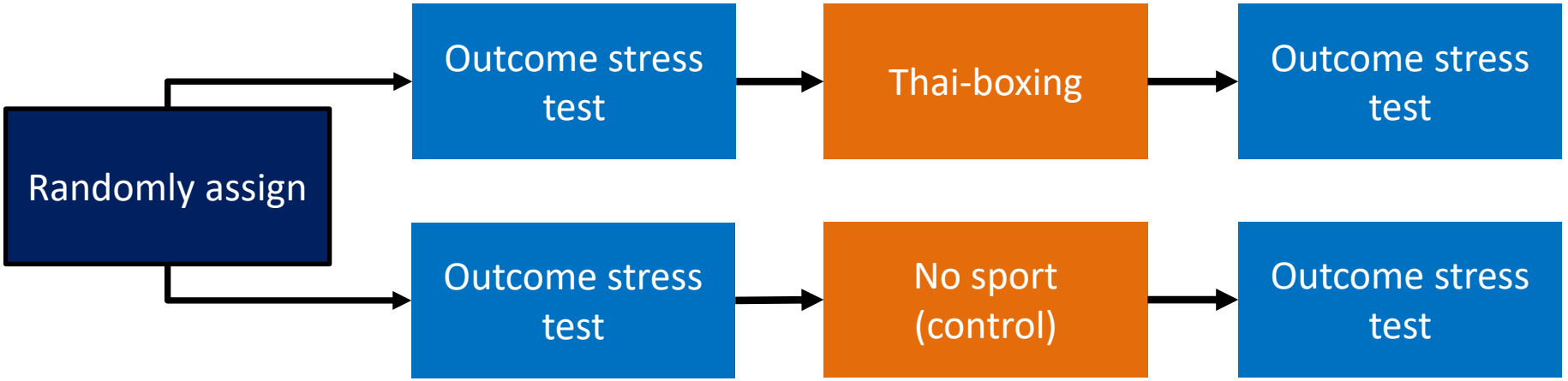
Chapter 10: independent-groups design (posttest only):



Experiments

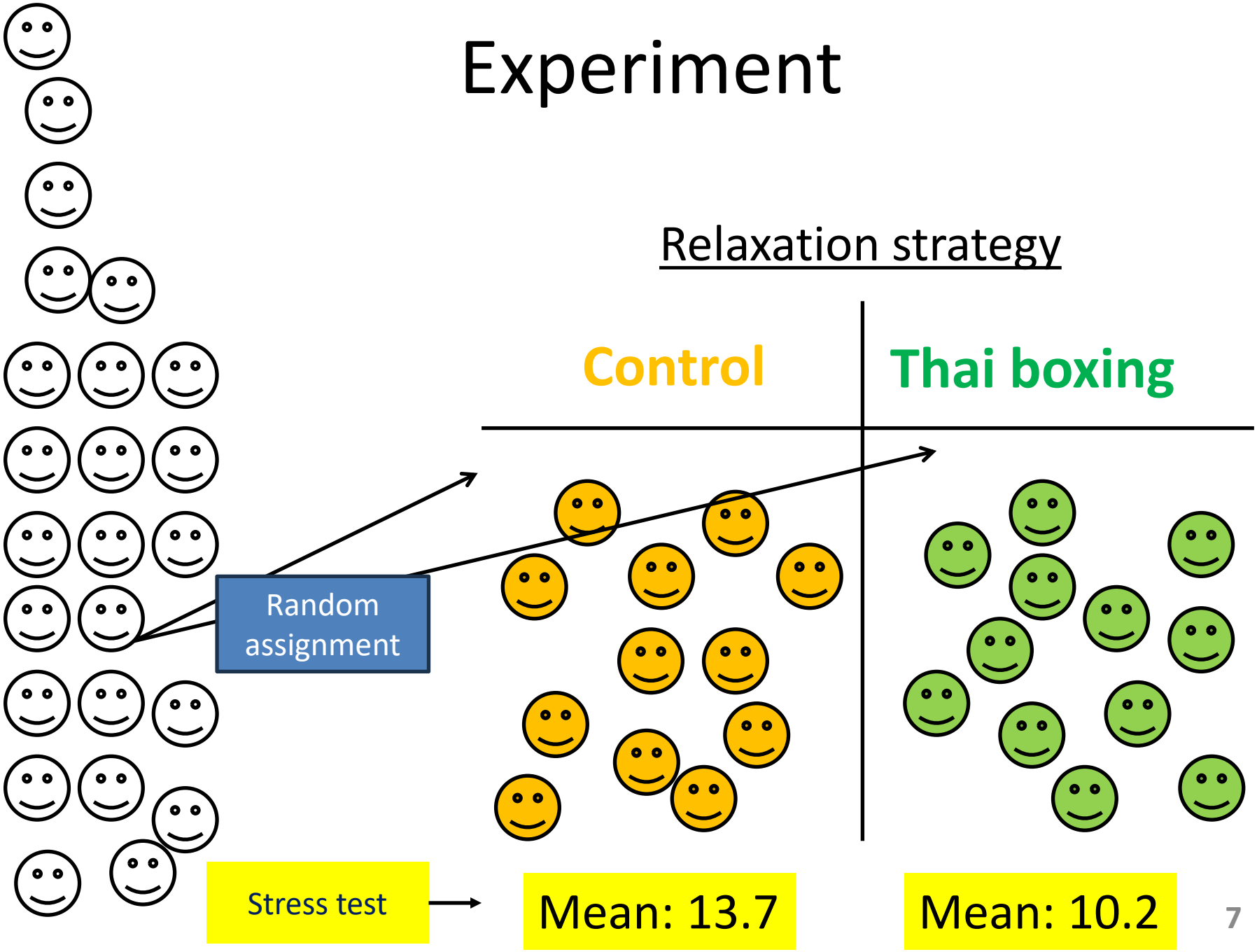
How to study whether Thai boxing is an effective stress relief strategy?

Chapter 10: pretest/posttest design:



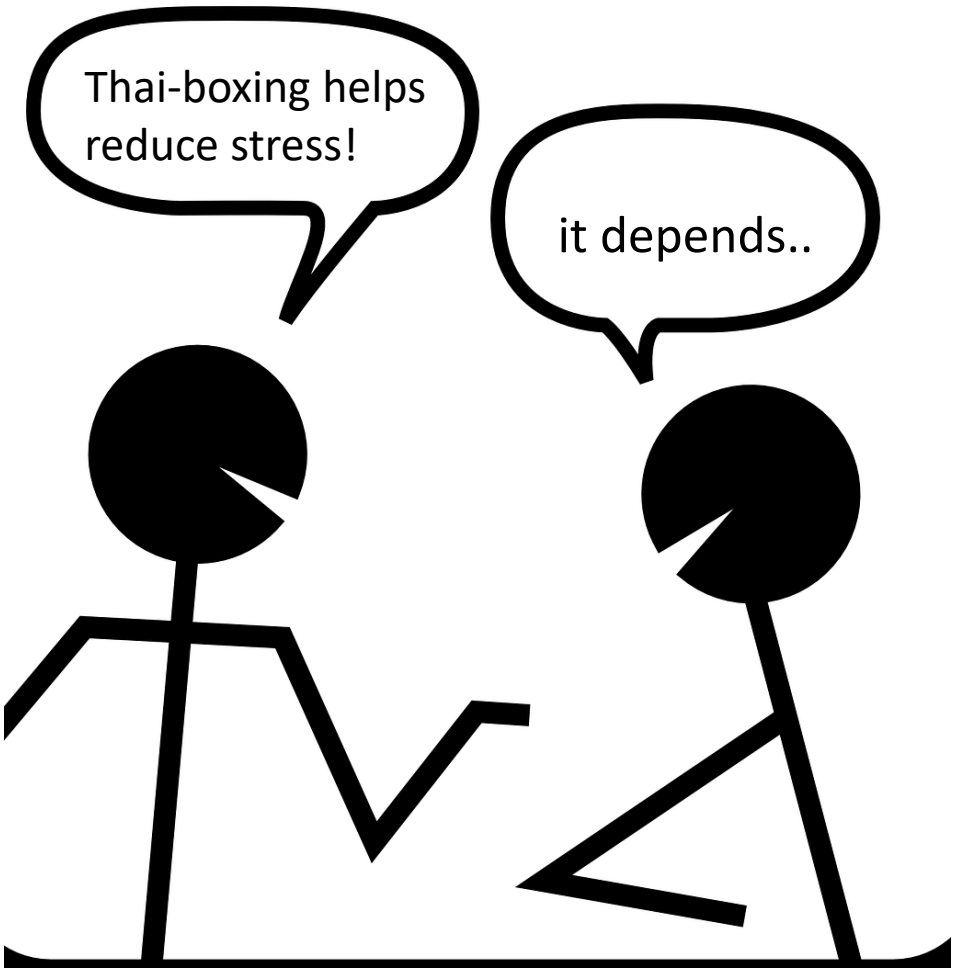
But let's for simplicity now stick with a posttest only design!

Experiment


Relaxation strategy
Control
Thai boxing
Random assignment
Stress test
Mean: 13.7
Mean: 10.2

Experiment

Suppose we found this difference in stress scores…



The (strength of the) effect might depend on….

- age?

- familiarity with boxing?

- cause of the stress?

- ….?

Picture source: pixabay.com

Experiment

Suppose we found this difference in stress scores…

If the groups do not systematically differ on these factors (which we try avoid by using random assignment), they are not confounders.

So we can say there is a “main-effect” of kickboxing on stress level.

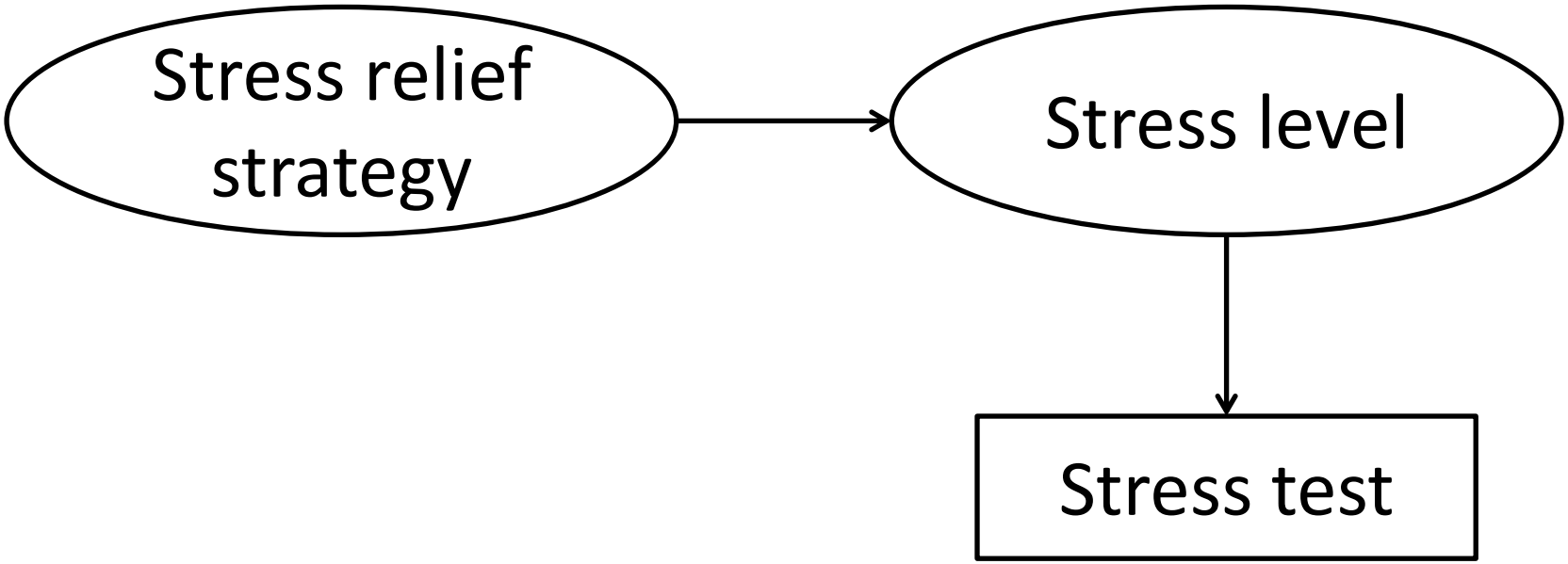
But, now we might be interested in whether the (strength of the) effect depends on other factors: interactions!

Main effect: is there a difference?
Interaction: is there a difference in differences?

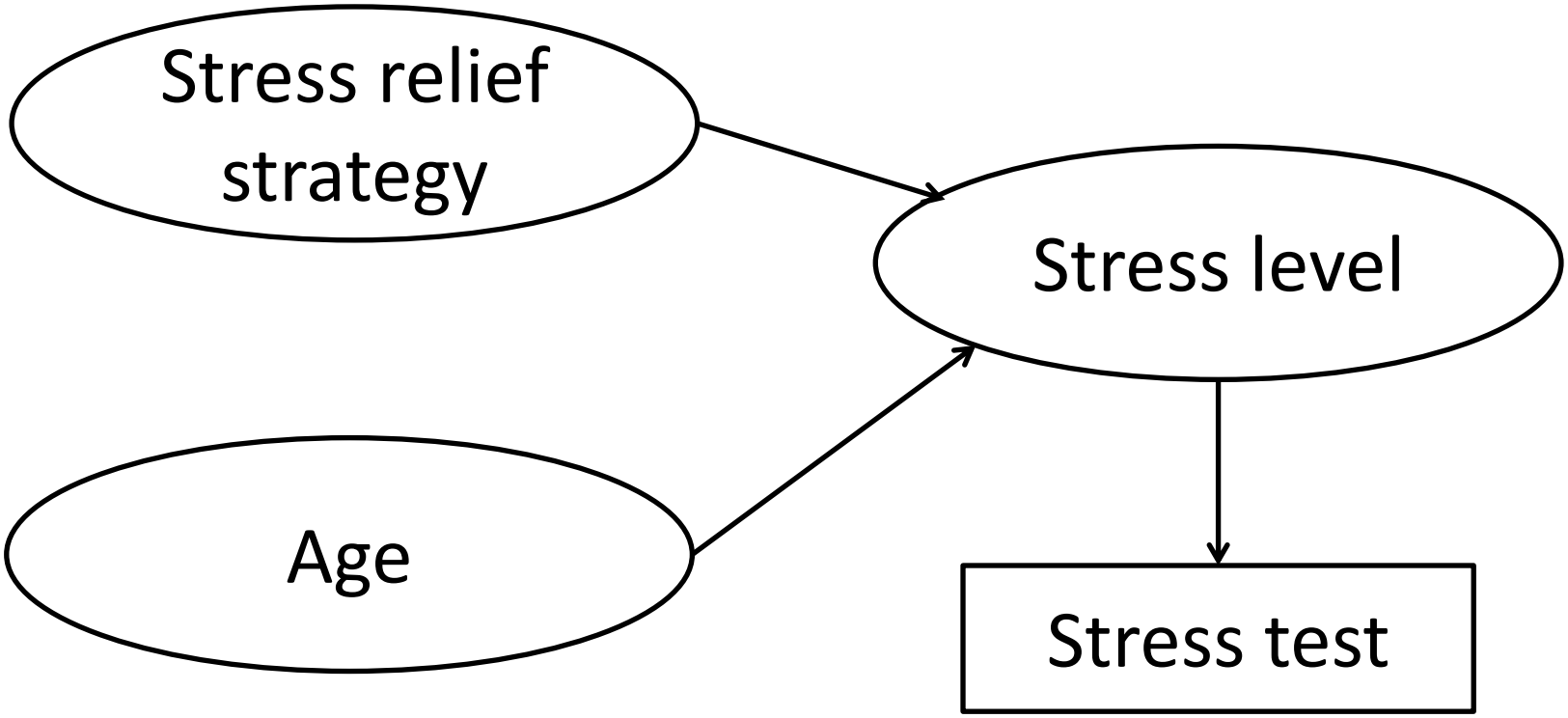
The (strength of the) effect might depend on….

- age?
- familiarity with boxing?
- cause of the stress?
- ….?

Independent-groups design with one independent variable



Independent-groups design with more than one independent variable (Ch12)



Today

Factorial design

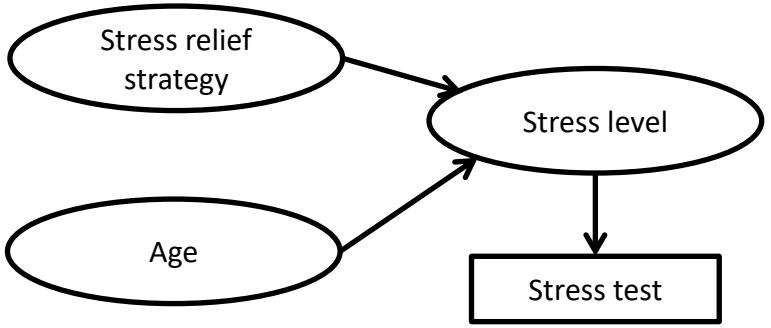
- **Two independent variables**
  - Adding more levels
- Three independent variables
- Factorial variations

Obscuring factors

- Not enough variability between groups
- Too much variability within groups

Factorial design: two or more independent variables

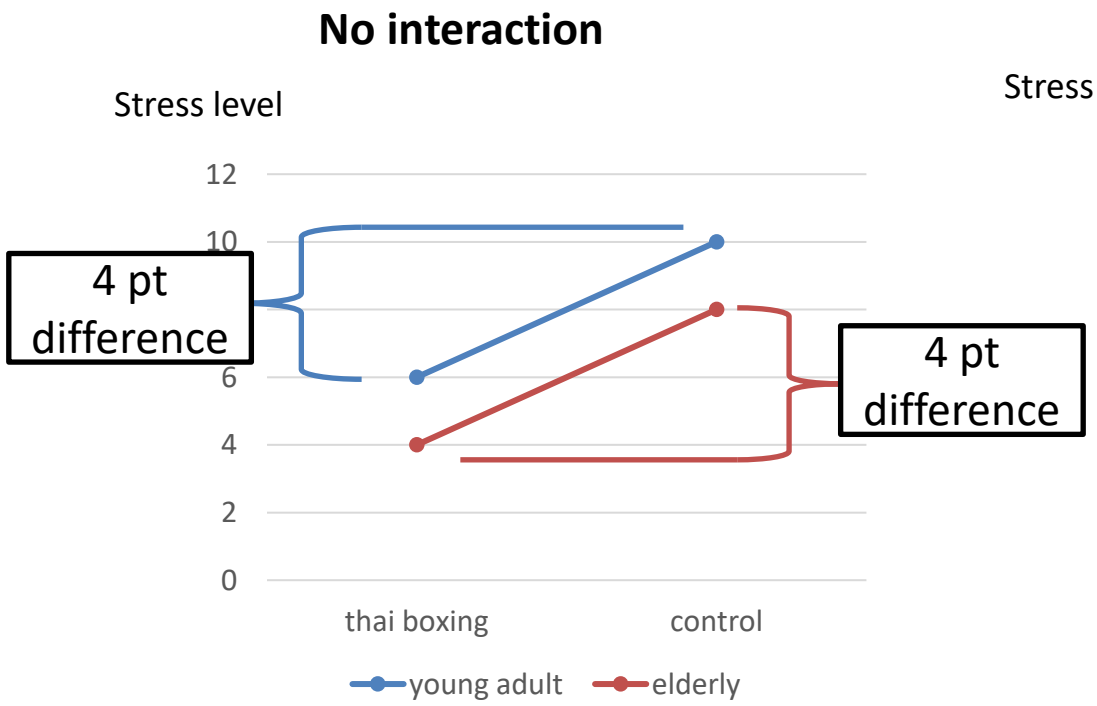
- Advantages:
  - More than one possible main effect
  - You can study *interaction effects*

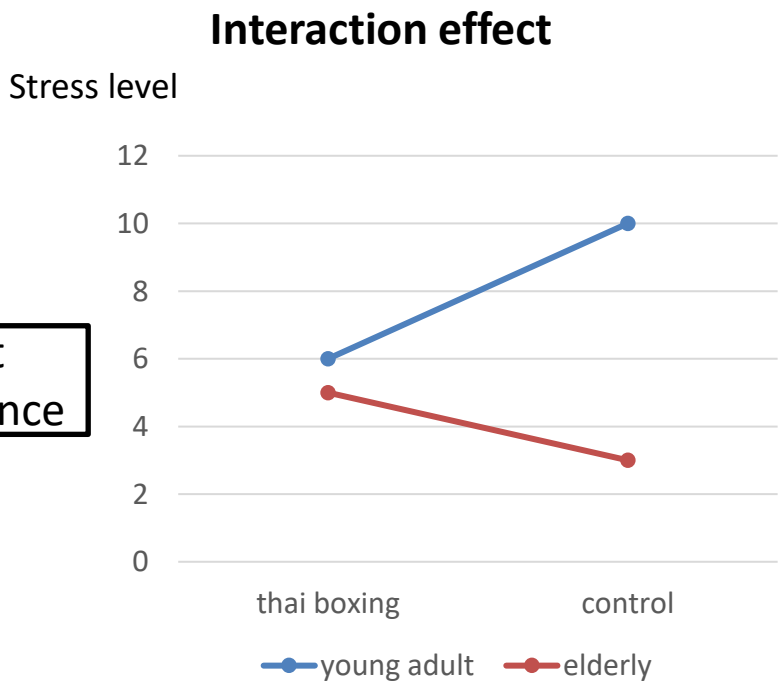


- Disadvantages:
  - You need more participants (at least twice) for the same accuracy

Interaction effects

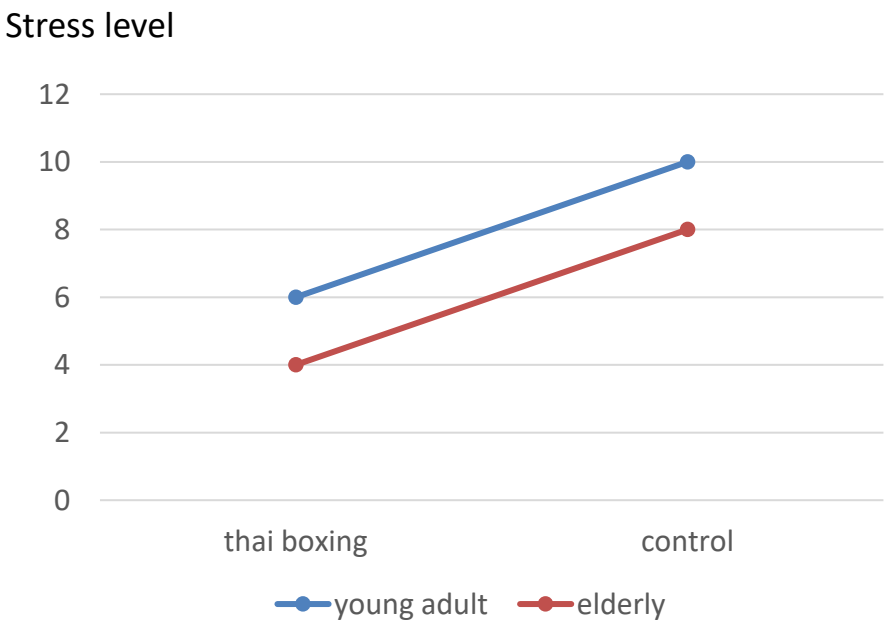
- The effect of one independent variable (ID) on the dependent variable (DV) depends on another variable

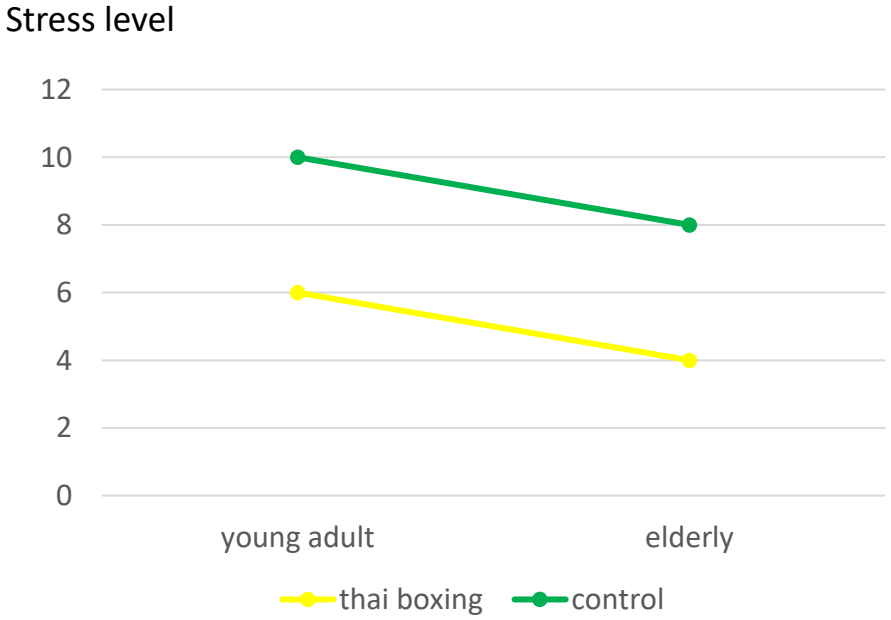




Interaction effects

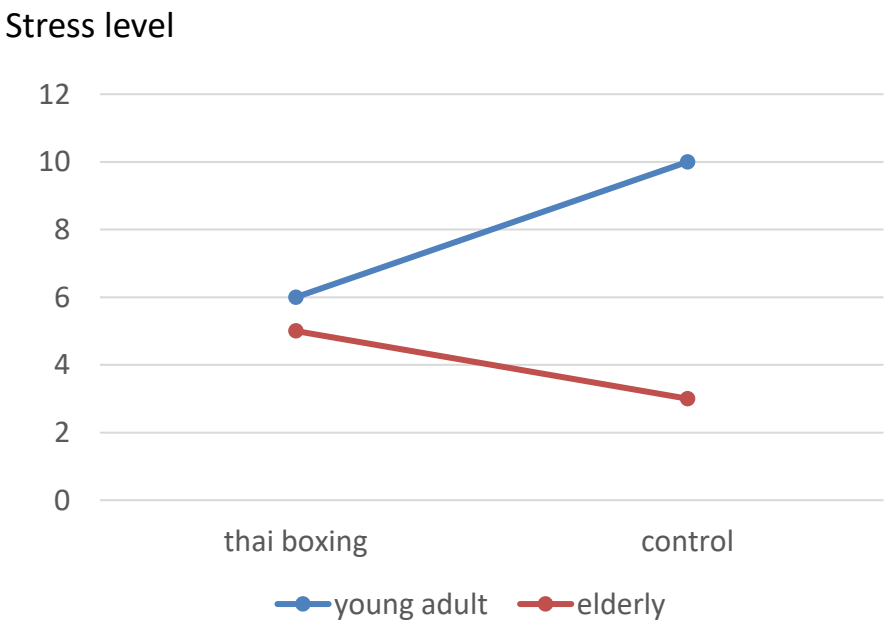
- Interaction effects are symmetric
  - **If the effect of ID1 on DP does *not* depend on ID2, then also the effect of ID2 on DP does *not* depend on ID1 (= no interaction)**
  - If the effect of ID1 on DP depends on ID2, then also the effect of ID2 on DP depends on ID1 (= interaction)

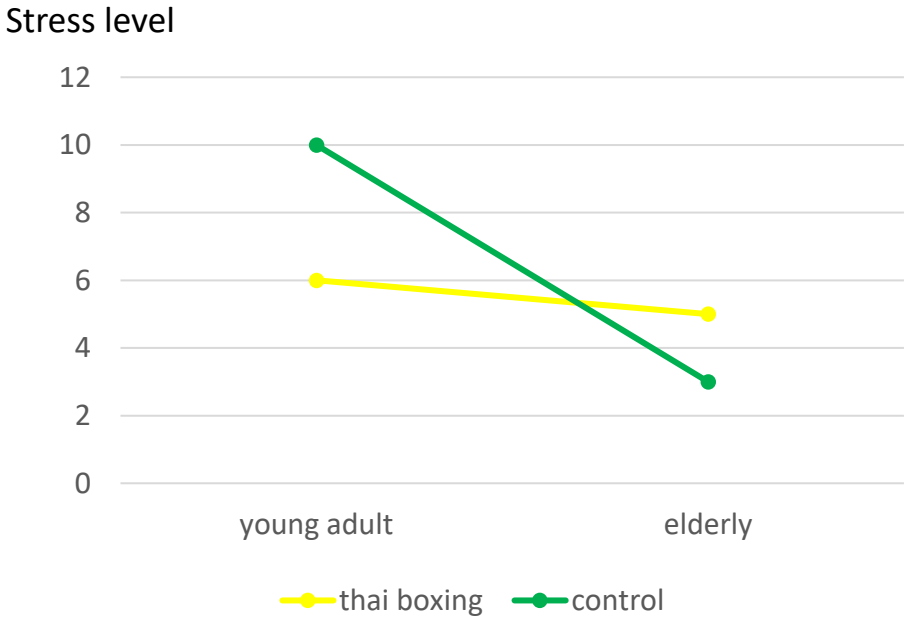




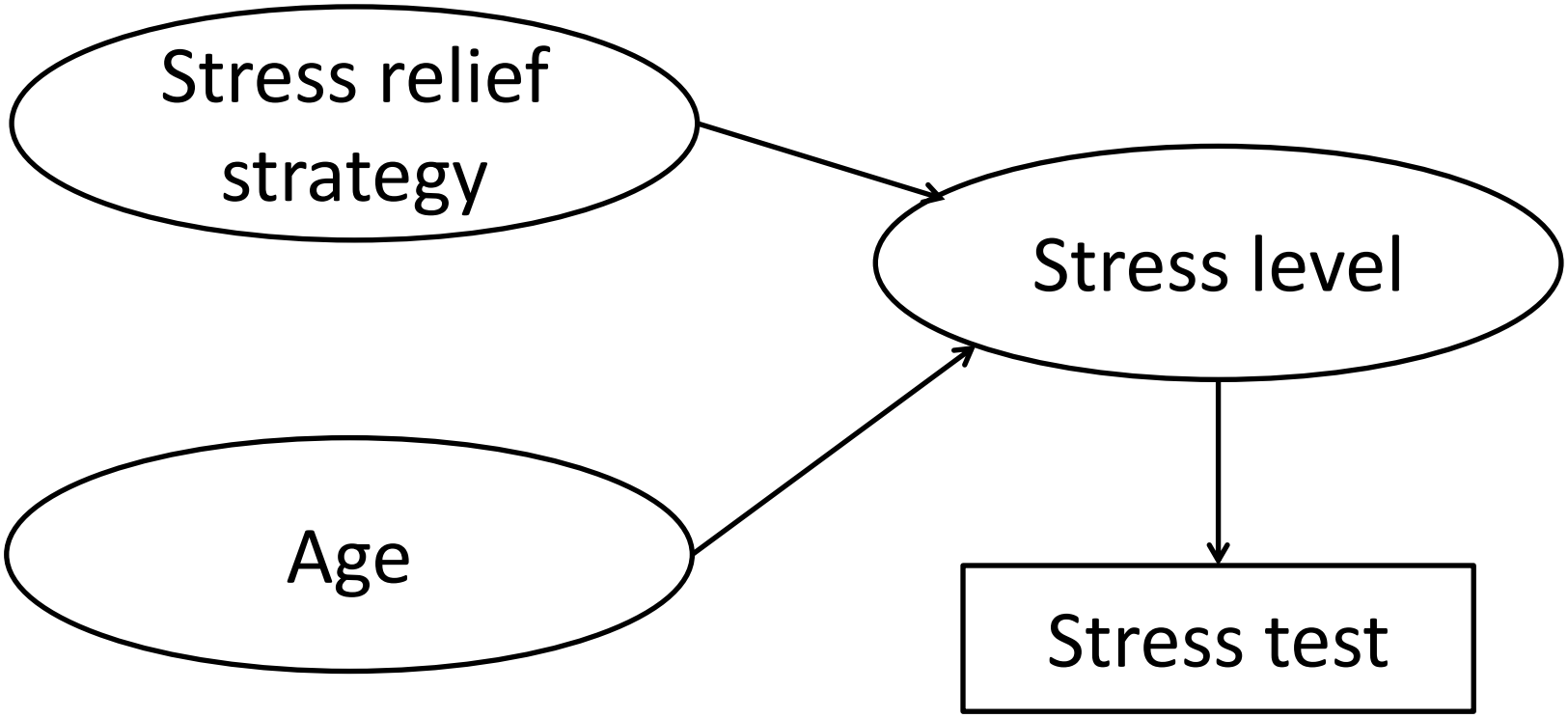
Interaction effects

- Interaction effects are symmetric
  - If the effect of ID1 on DP does *not* depend on ID2, then also the effect of ID2 on DP does *not* depend on ID1 (= no interaction)
  - **If the effect of ID1 on DP depends on ID2, then also the effect of ID2 on DP depends on ID1 (= interaction)**

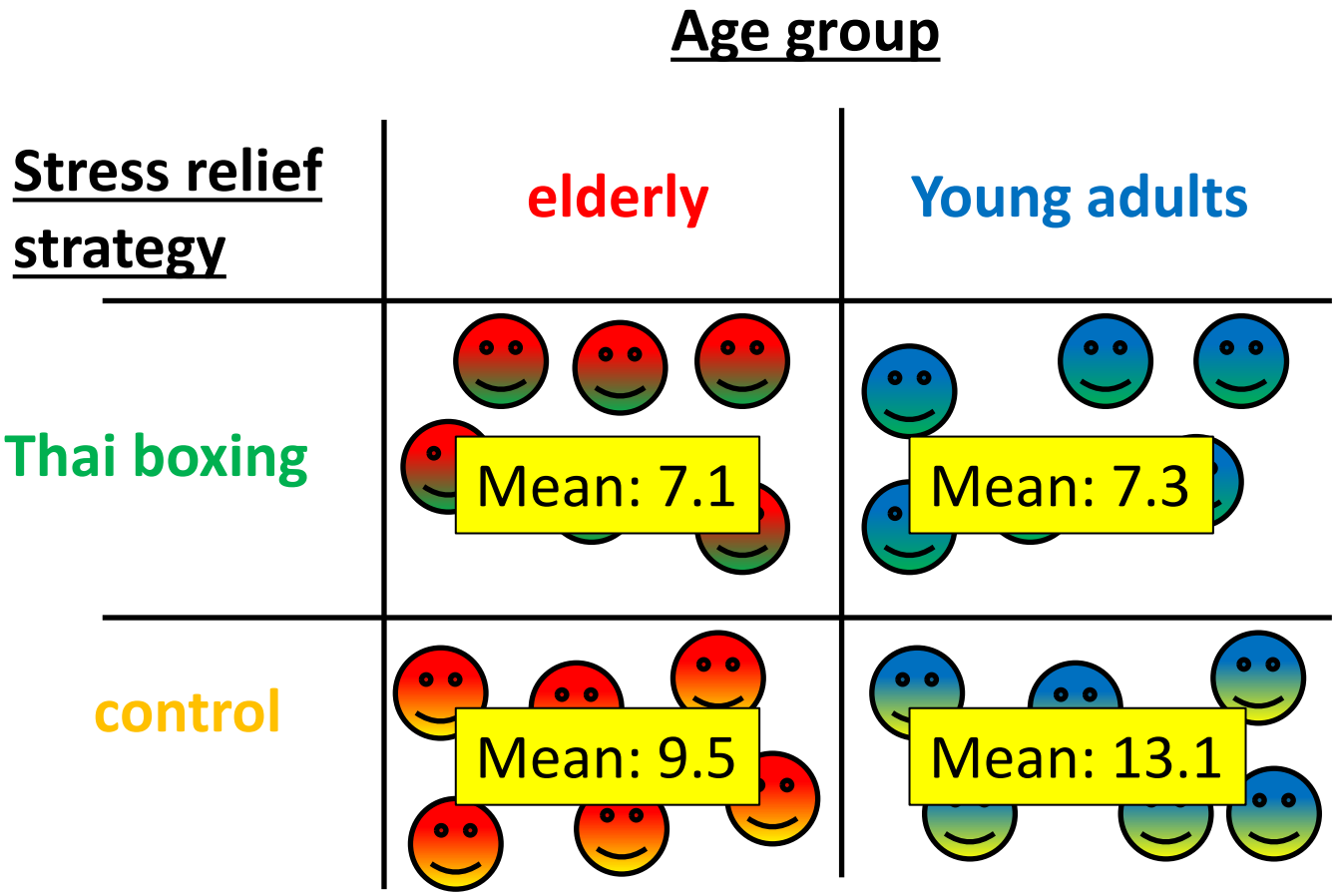




Independent-groups design with more than one independent variable (Ch12)



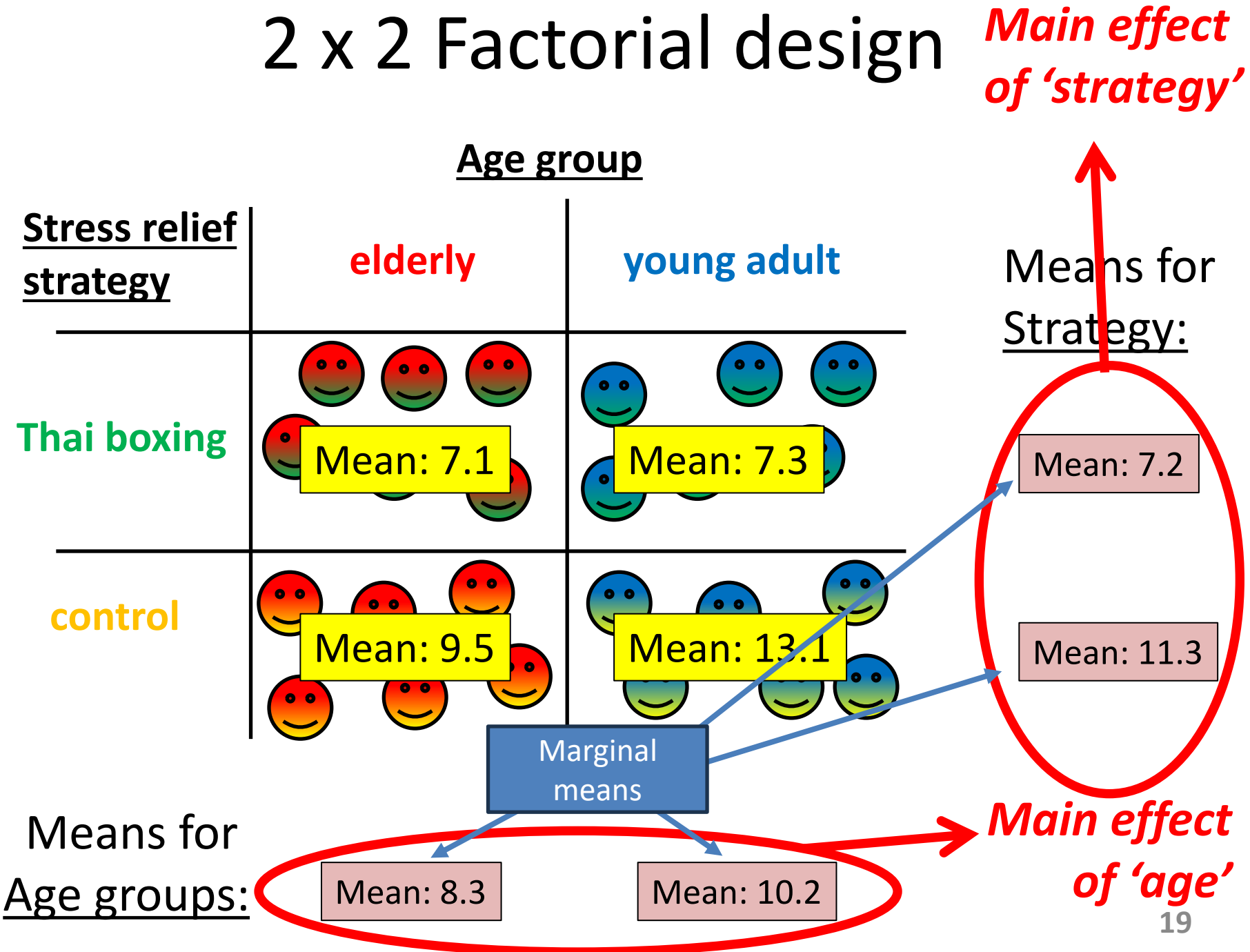
2 x 2 Factorial design



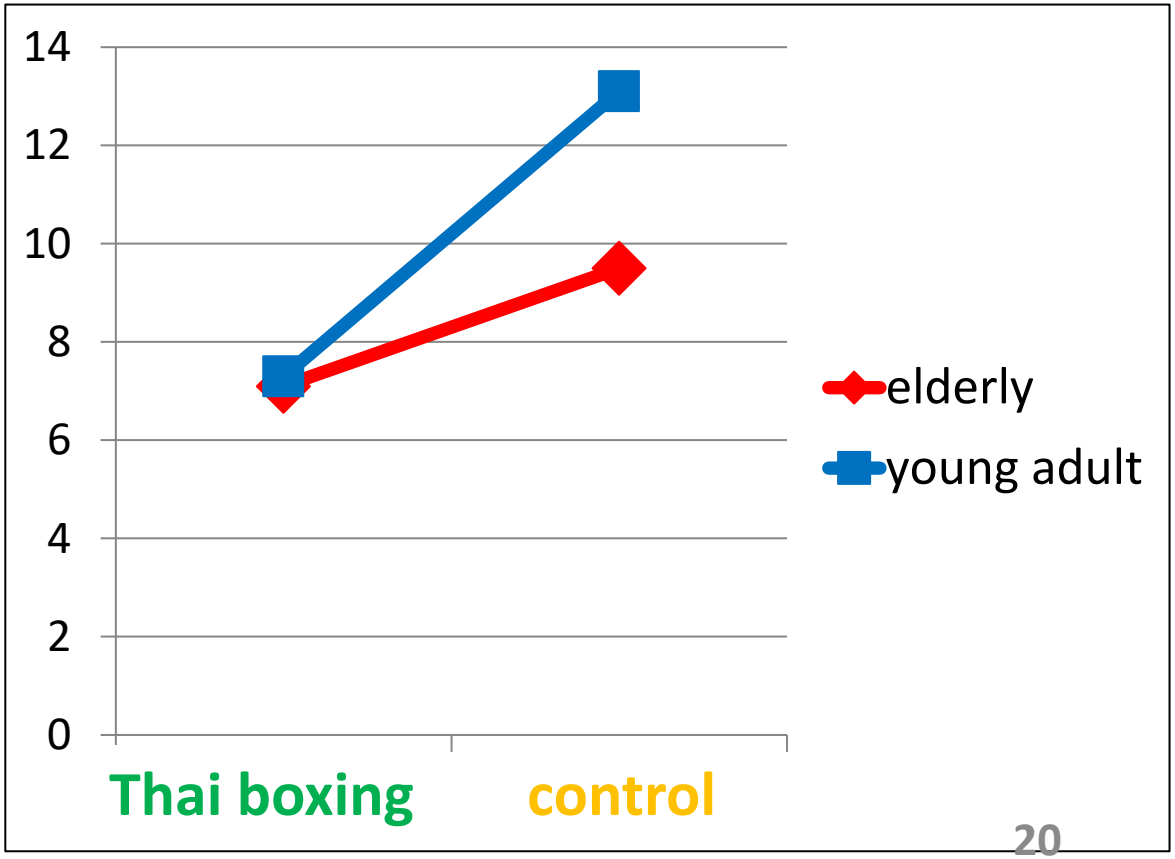
2 x 2 because 2 independent variables:

- strategy (2 *levels*: control- thai boxing);
- age group (2 *levels*: young adult - elderly)

2 x 2 Factorial design



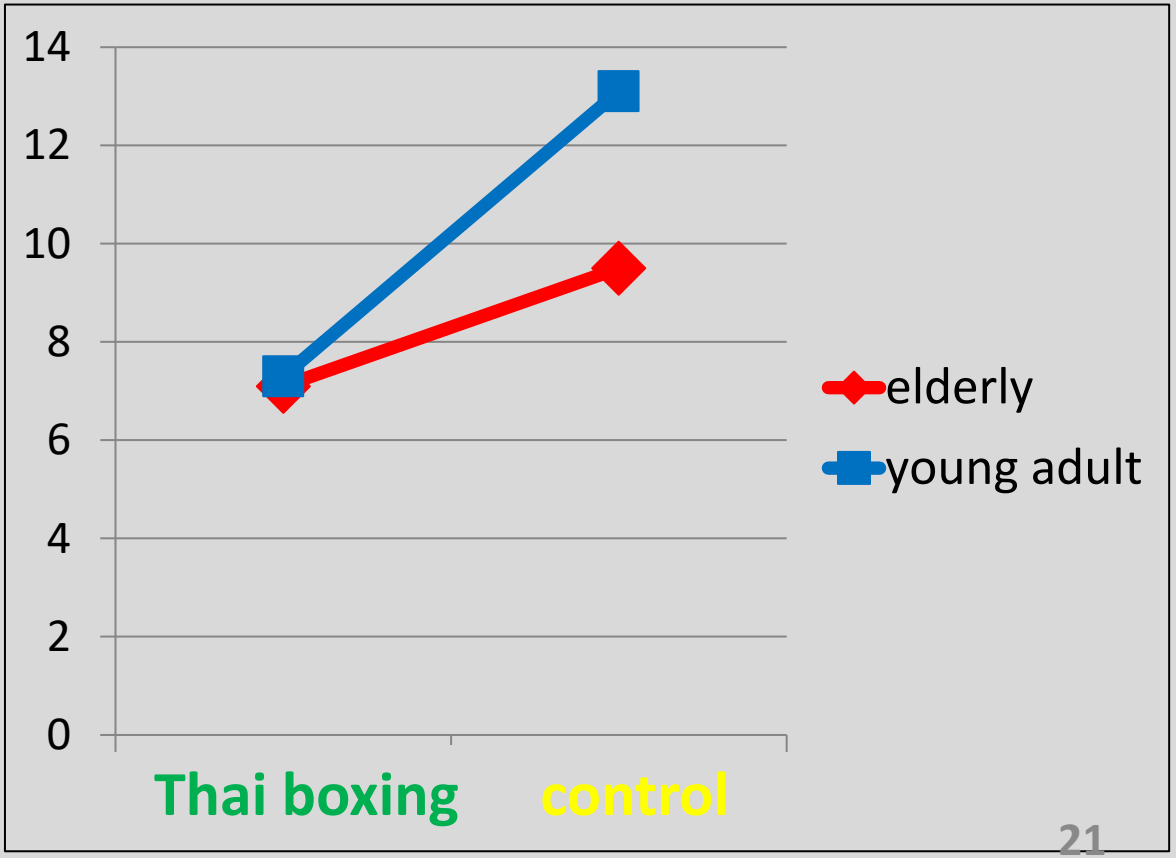
Strategy	Age: elderly	Age: Young adults
Thai boxing	7.1	7.3
No sport	9.5	13.1



Practice

(also in the weekly assignment)

Main effect for age?



Practice

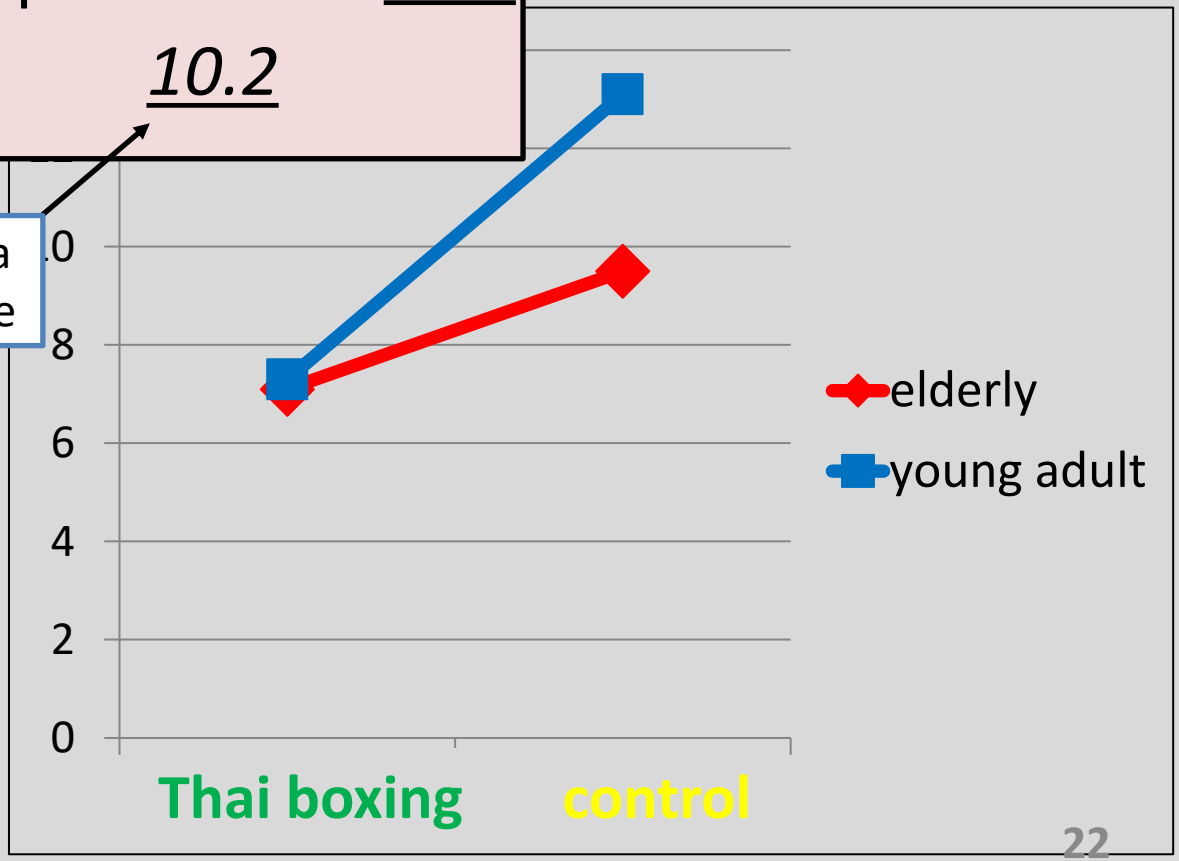
nment)

Strategy	Age: elderly	Age: Young adult	
Thai boxing	7.1	7.3	7.2
control	9.5	13.1	11.3
	8.3	10.2	

Calculate marginal means

Main effect for age? → YES

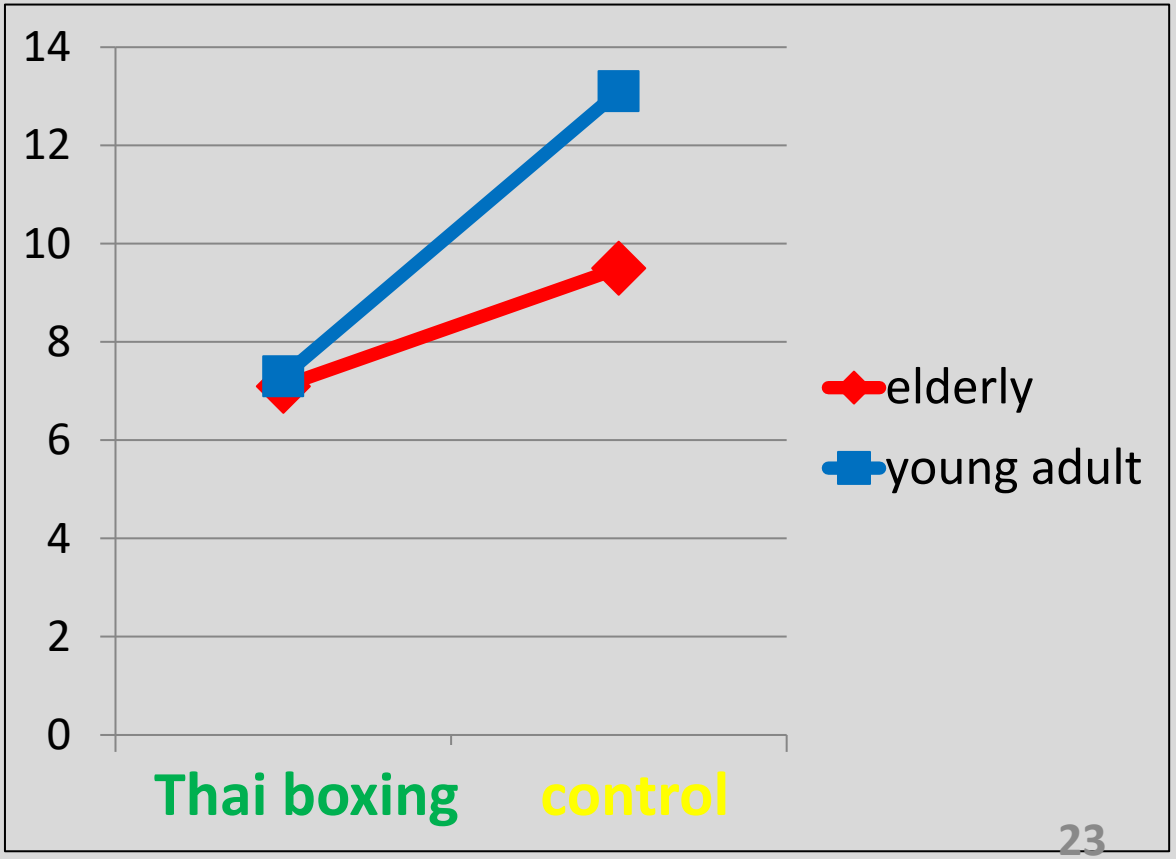
There is a difference



Practice

(also in the weekly assignment)

Main effect for strategy?



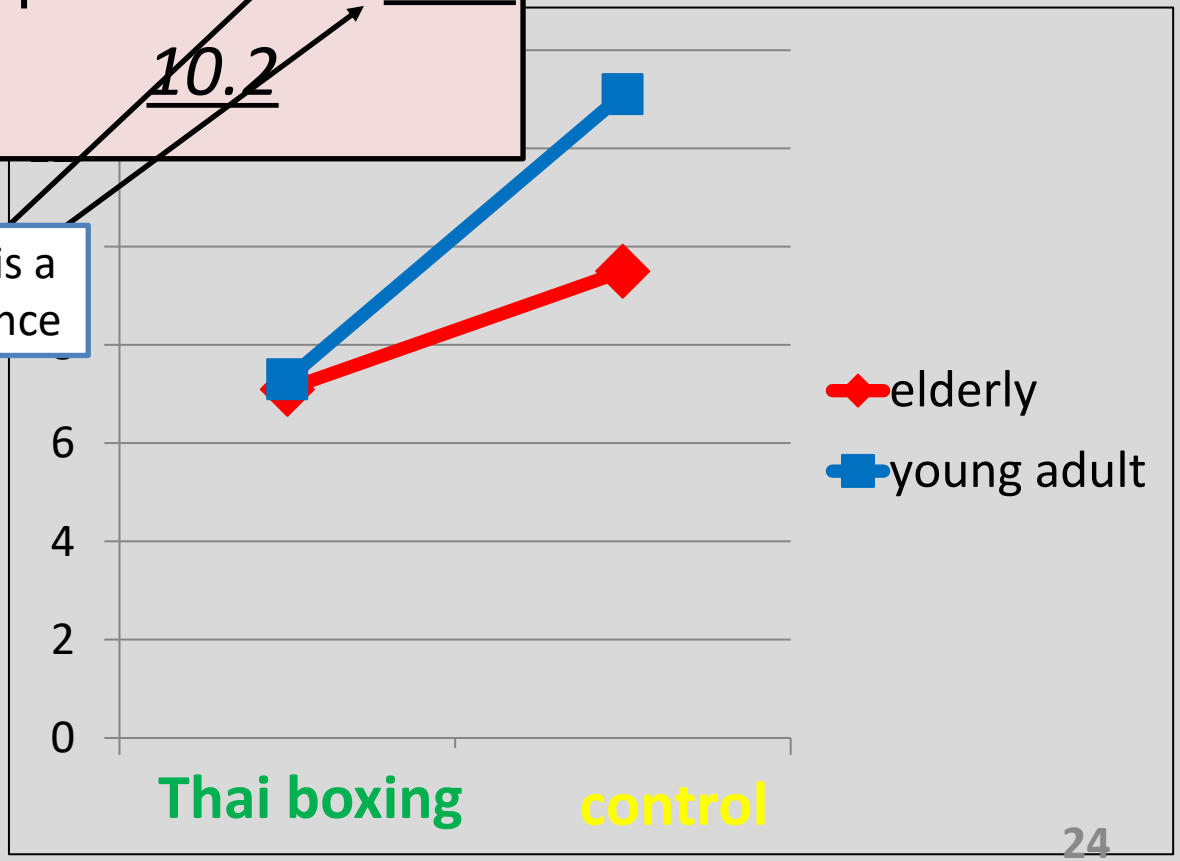
Practice

nment)

Strategy	Age: elderly	Age: Youg adult	
Thai boxing	7.1	7.3	7.2
control	9.5	13.1	11.3
	8.3	10.2	

Main effect for strategy? → YES

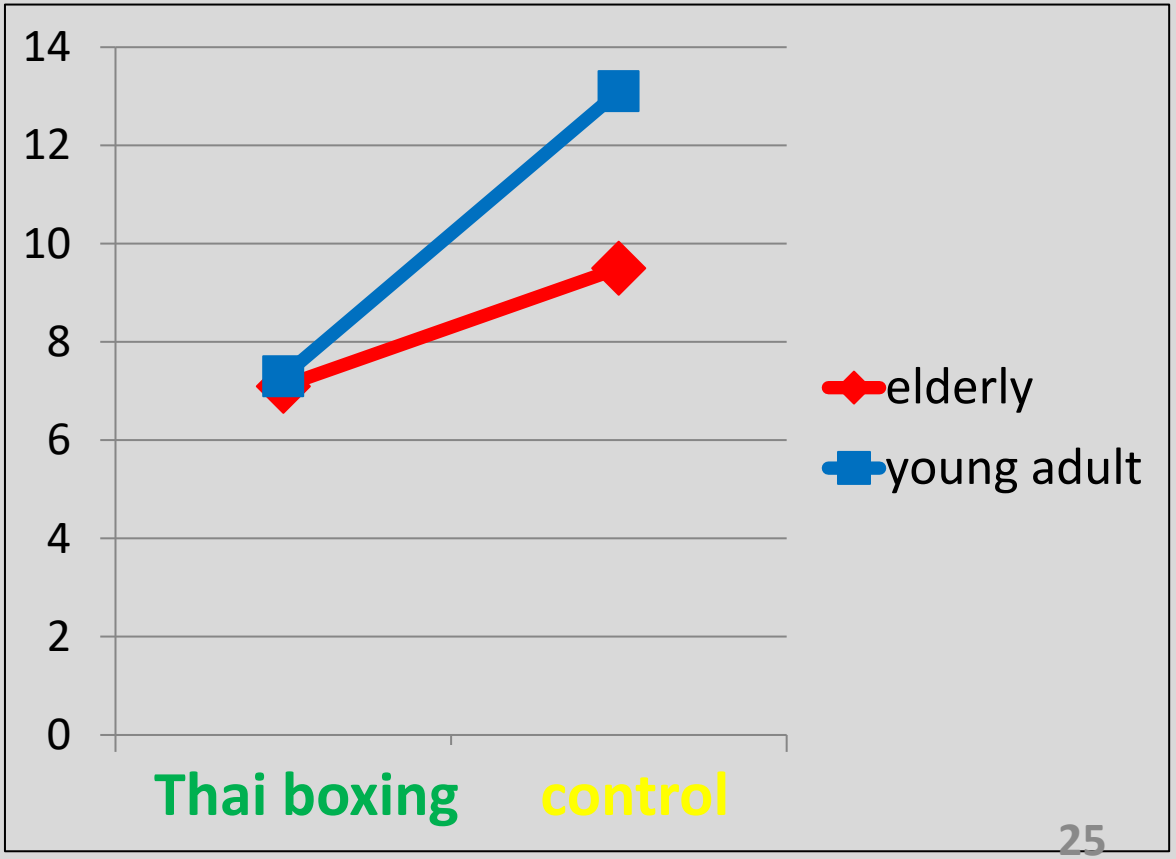
There is a difference



Practice

(also in the weekly assignment)

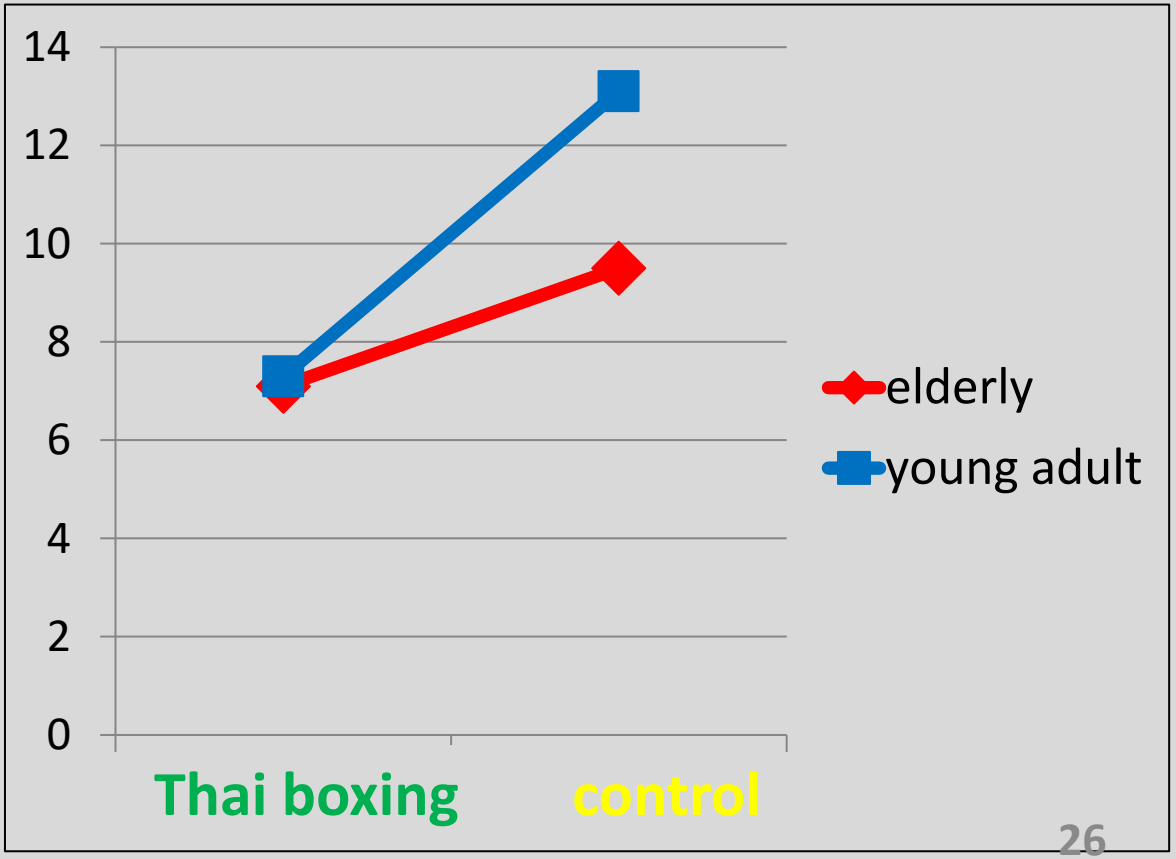
Interaction effect?



Practice

(also in the weekly assignment)

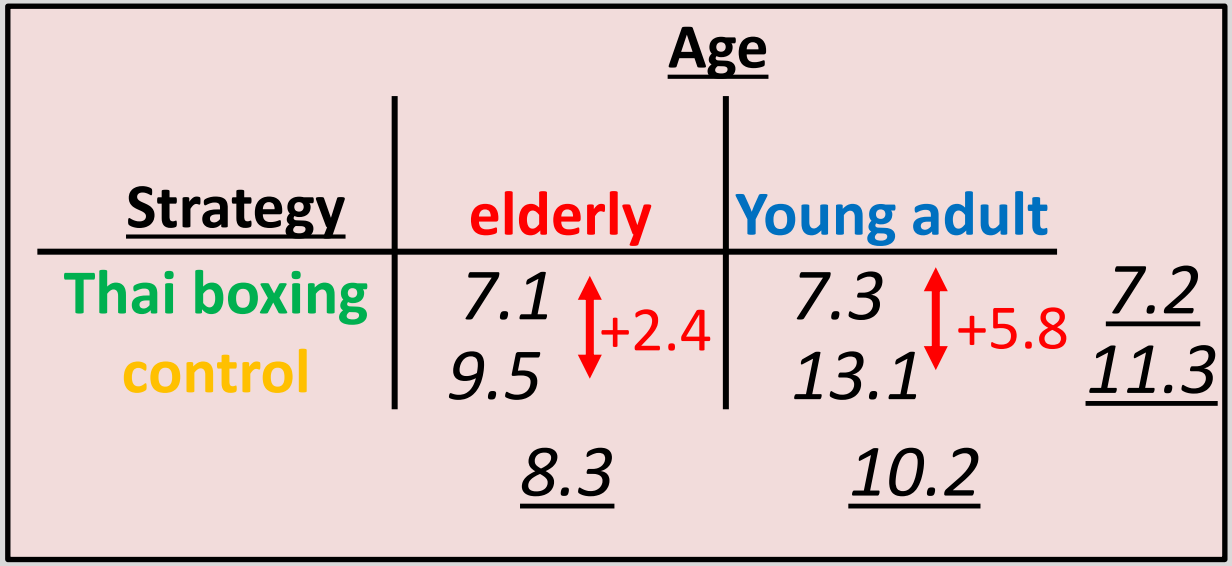
Interaction Effect?
→ YES



Trick:
If the lines are parallel, there is no interaction

Practice

(also in the weekly assignment)



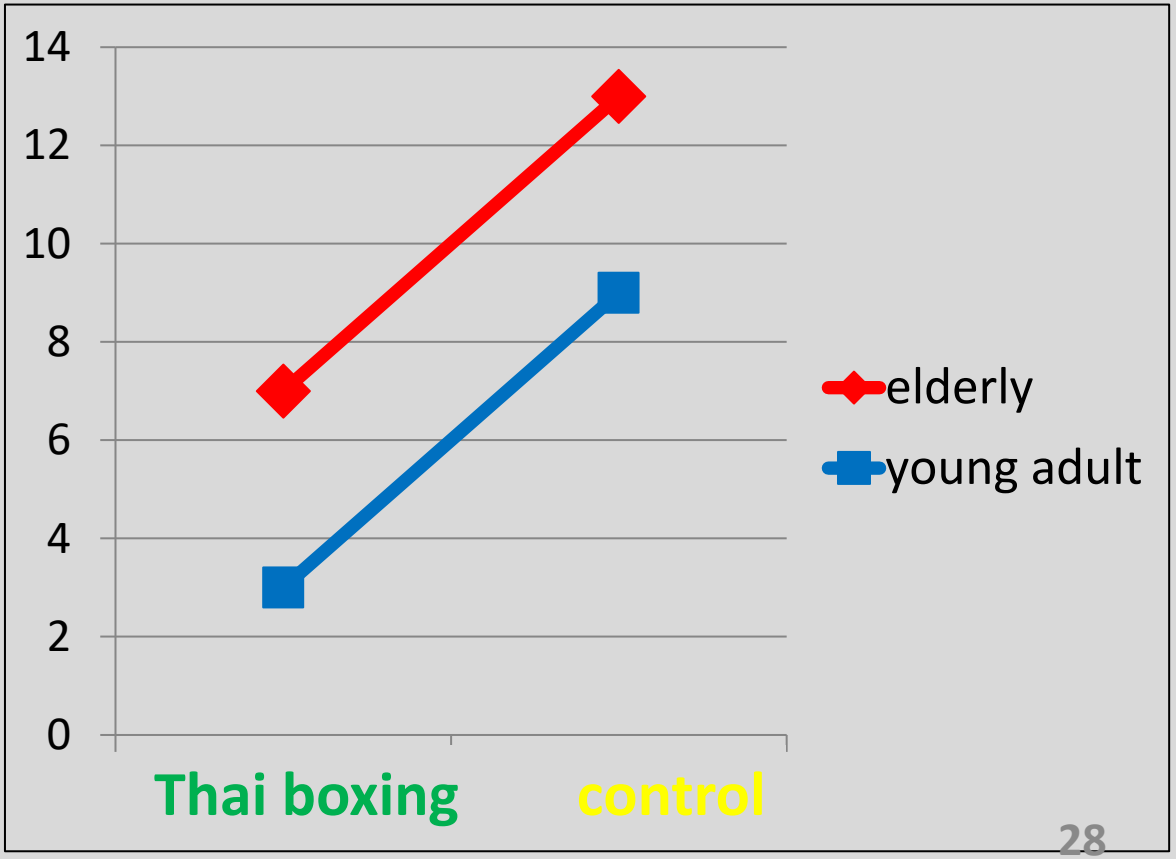
Interaction Effect? → YES

Trick:
If there is a difference in differences, there is an interaction

Practice

(also in the weekly assignment)

Main effects?

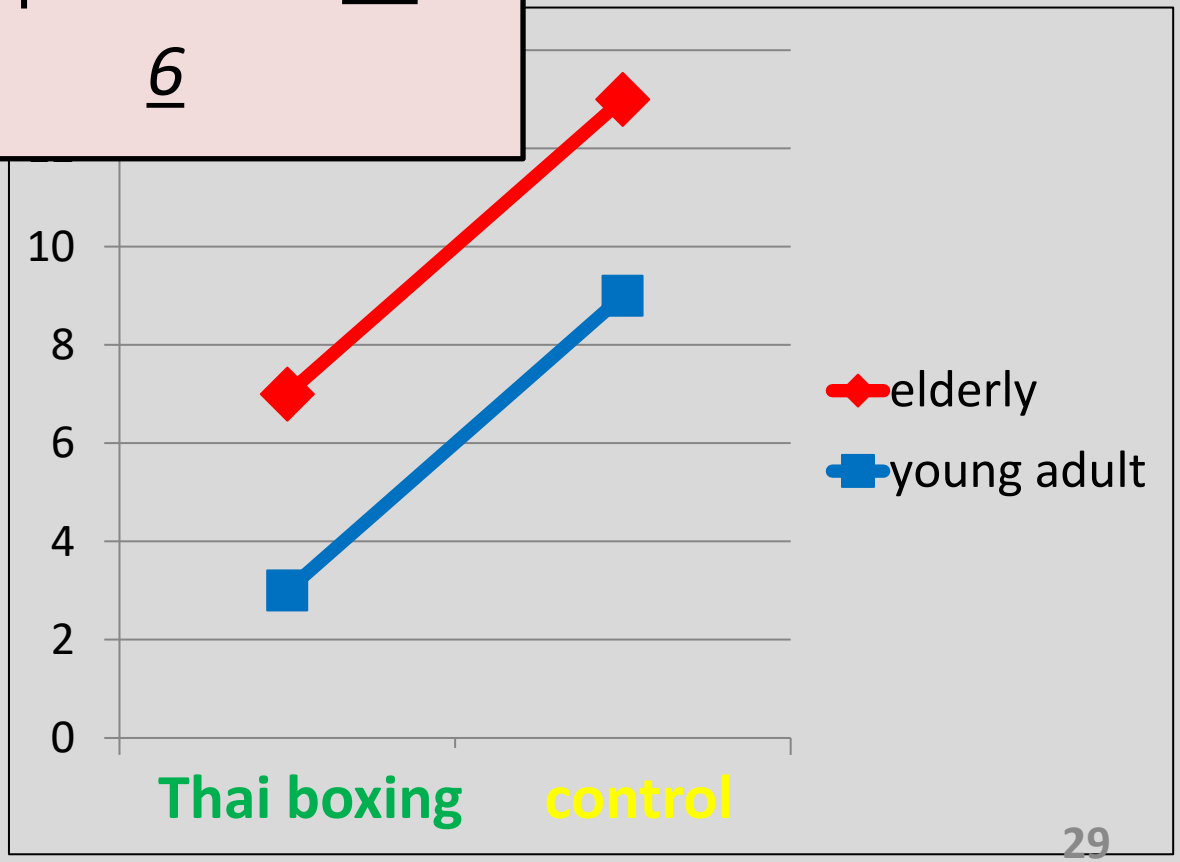


Practice

gnment)

Strategy	Age: elderly	Age: Young adult	
Thai boxing	7	3	5
control	13	9	11
	10	6	

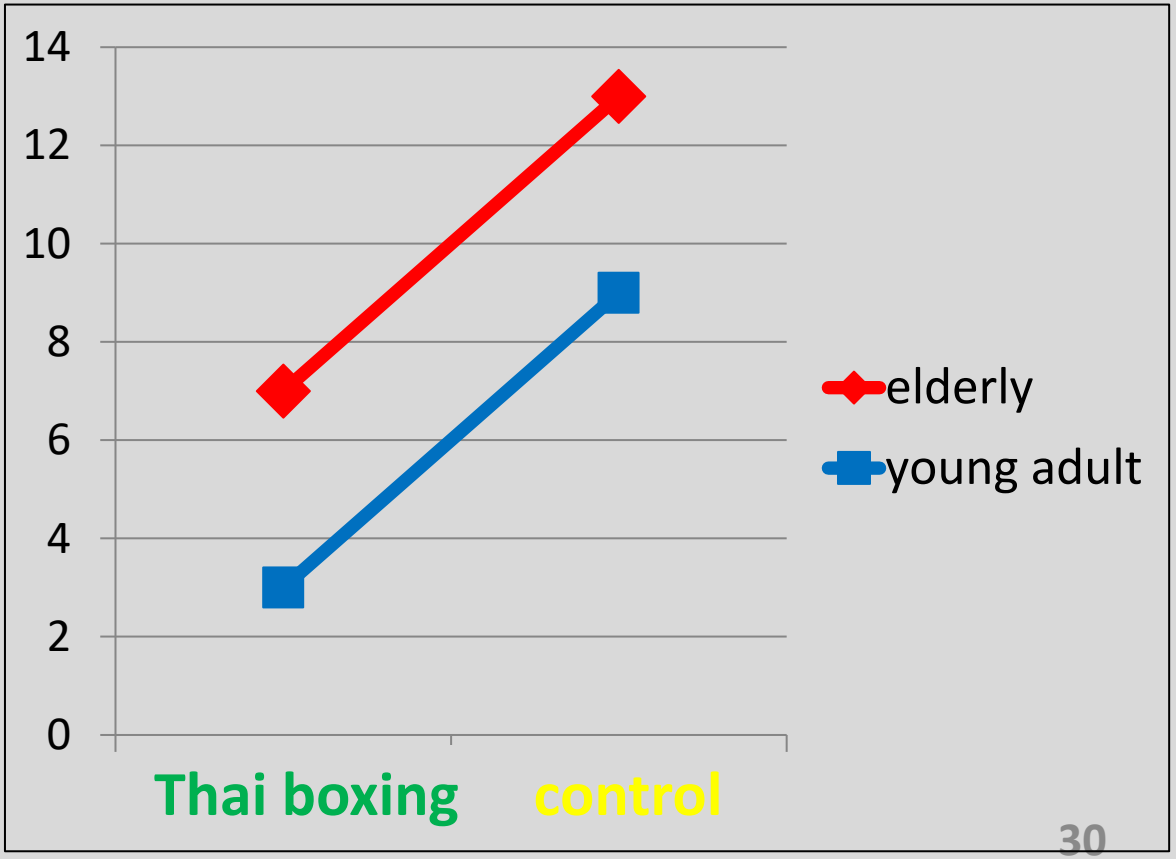
Main effects?
→ YES



Practice

(also in the weekly assignment)

Interaction Effect?



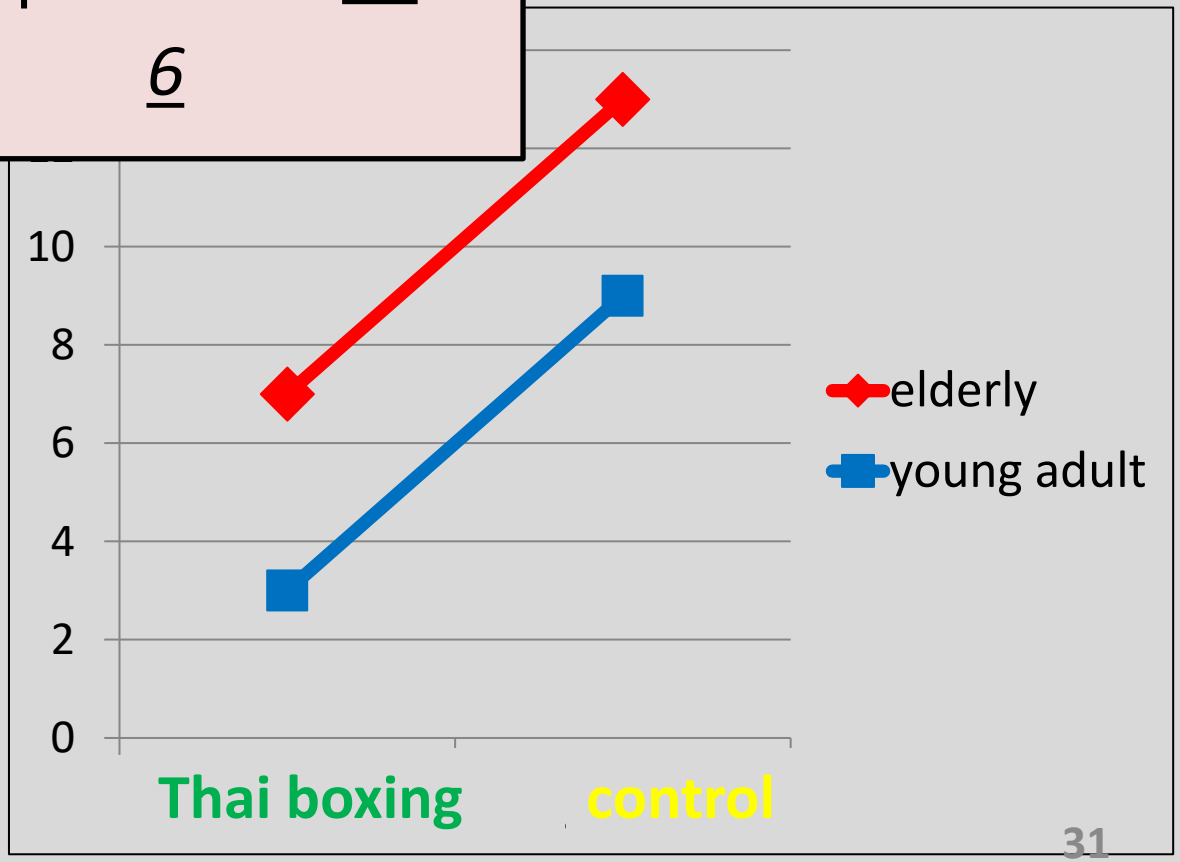
Practice

gnment)

Strategy	Age: elderly	Age: Young adult	
Thai boxing	7	3	5
control	13	9	11
	10	6	

+6 +6

Interaction Effect?
→ NO

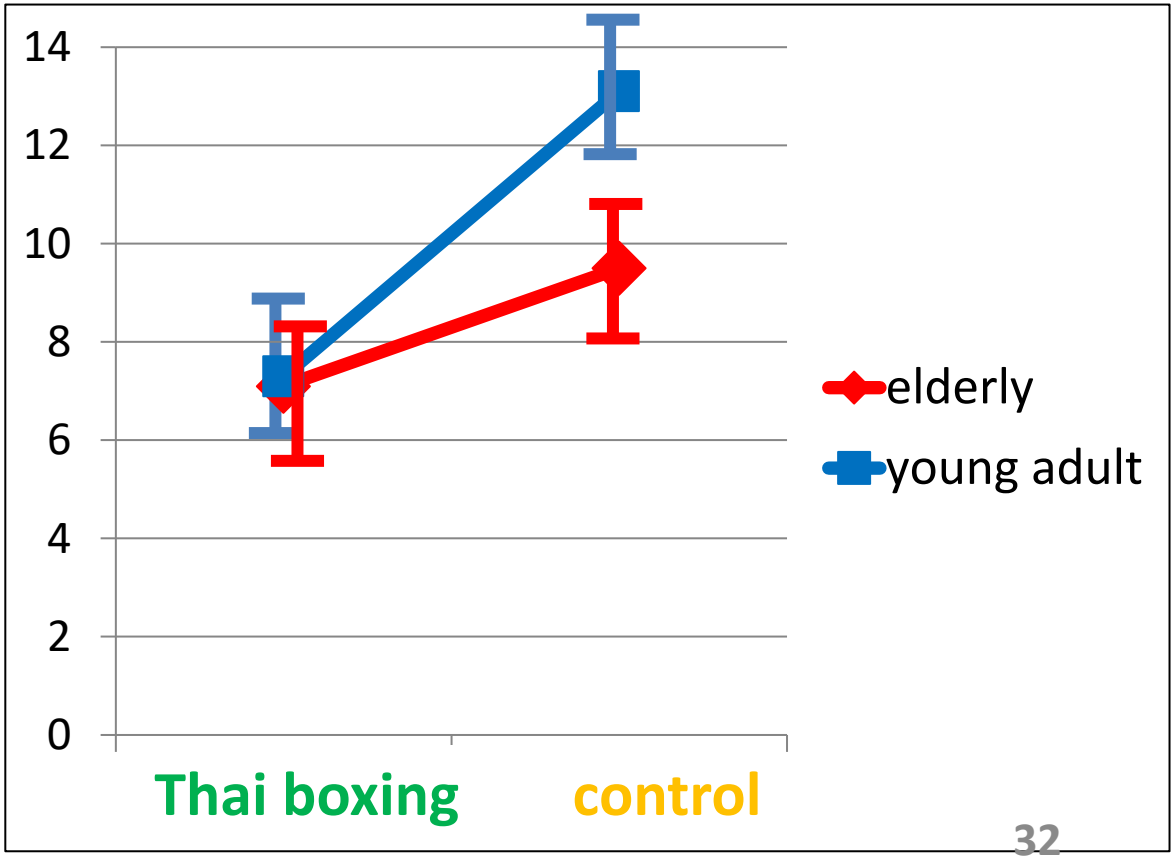


In these cases we drew conclusions about main and interaction effects from tables and graphs without considering the uncertainty of these estimates:

"ignore sampling variability and only consider point estimates"

In practice, we would account for this uncertainty and consider whether the differences are 'significant'.

	Age	
Strategy	elderly	Young adults
Thai boxing	7.1	7.3
No sport	9.5	13.1



Today

Factorial design

- Two independent variables
  - **Adding more levels**
- Three independent variables
- Factorial variations

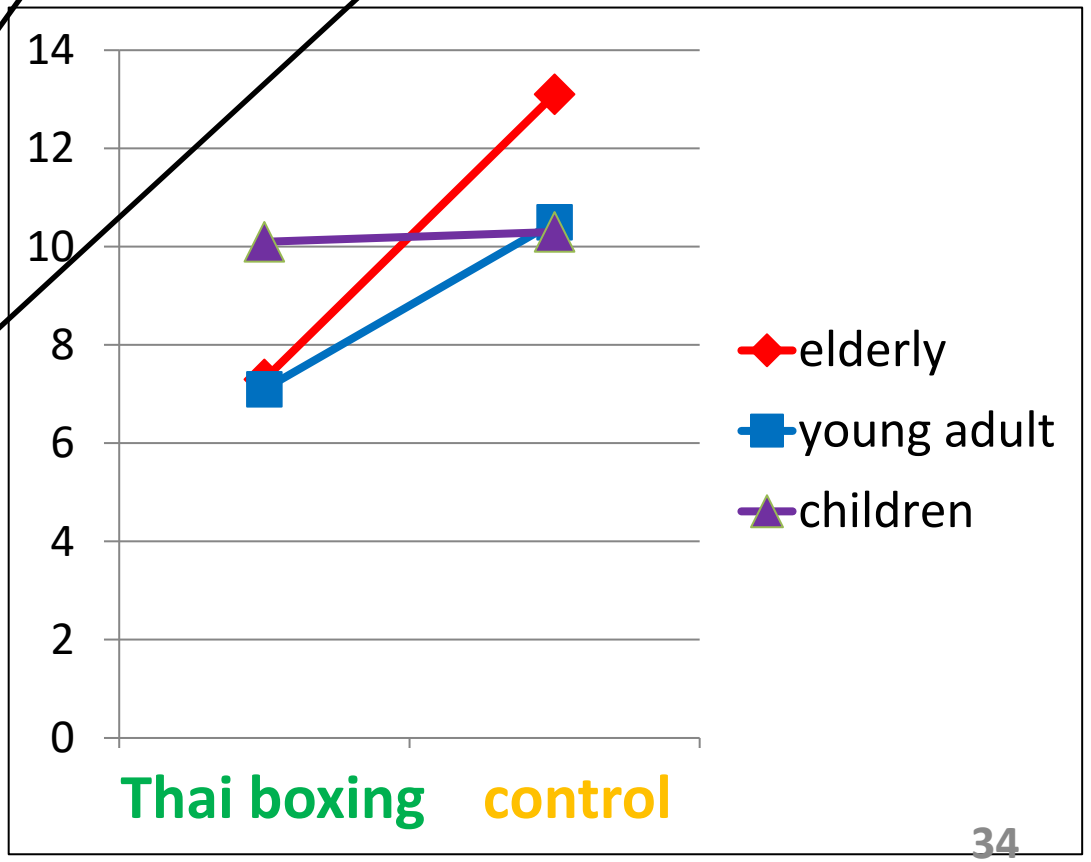
Obscuring factors

- Not enough variability between groups
- Too much variability within groups

	elderly	Young adult	children	*Mean strategy*
Thai boxing	7.3	7.1	10.1	*8.2*
control	13.1	10.5	10.3	*11.3*
mean age	*10.2*	*8.8*	*10.2*	

Extension:
Number of levels

E.g.,: 3 x 2 design

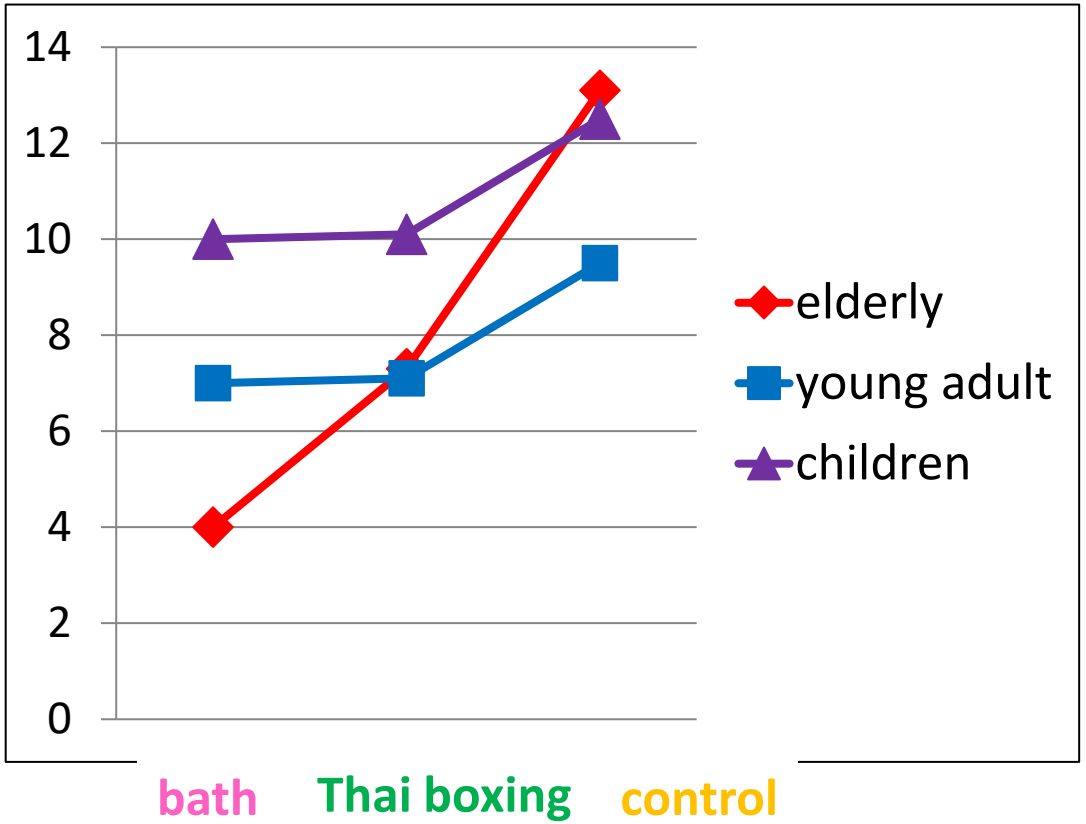


There is a main effect of age because there is a difference

	elderly	young adult	children	*mean strategy*
taking a bath	4.0	7.0	10.0	*7*
	+3.3	+0.1	+0.1	
Thai boxing	7.3	7.1	10.1	*8.17*
	+5.8	+2.4	+2.4	
control	13.1	9.5	12.5	*11.03*
mean age	*8.1*	*7.9*	*9.3*	

Extension:
Number of levels

E.g.,: 3 x 3 design



Today

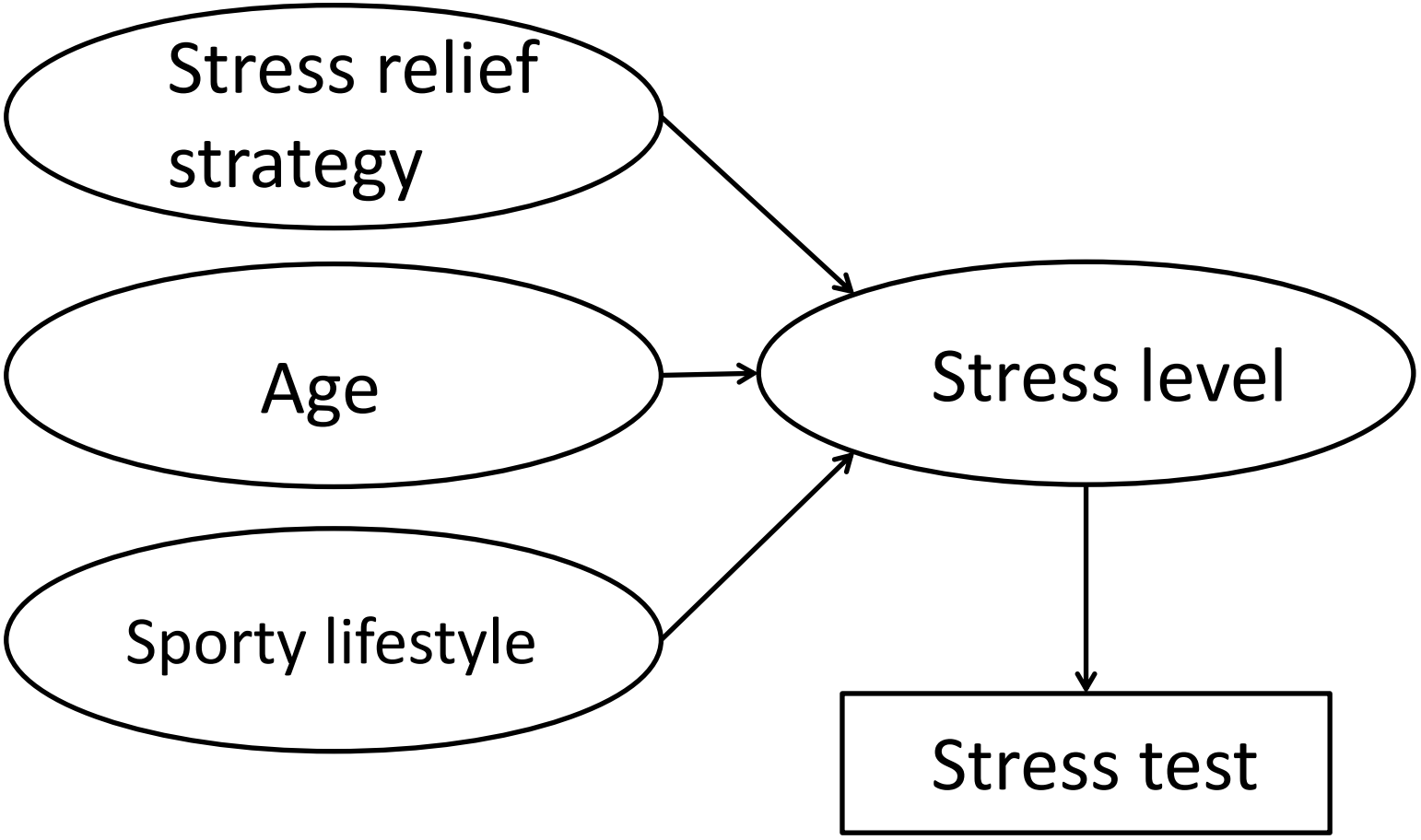
Extension on standard between-subjects design

- Two independent variables
  - Adding more levels
- **Three independent variables**
- Factorial variations

Obscuring factors

- Not enough variability between groups
- Too much variability within groups

Factorial design with three independent variables



2 x 2 x 2 factorial design

Lifestyle

sporty

	Age	
Status	**elderly**	**young adult**
Thai boxing	?	?
control	?	?

not sporty

	Age	
Strategy	**elderly**	**young adult**
Thai boxing	?	?
control	?	?



Source: pixabay

Interaction?

Lifestyle

sporty

Strategy	Age: elderly	Age: young adult
Thai boxing	?	?
control	?	?

not sporty

Strategy	Age: elderly	Age: young adult
Thai boxing	?	?
control	?	?

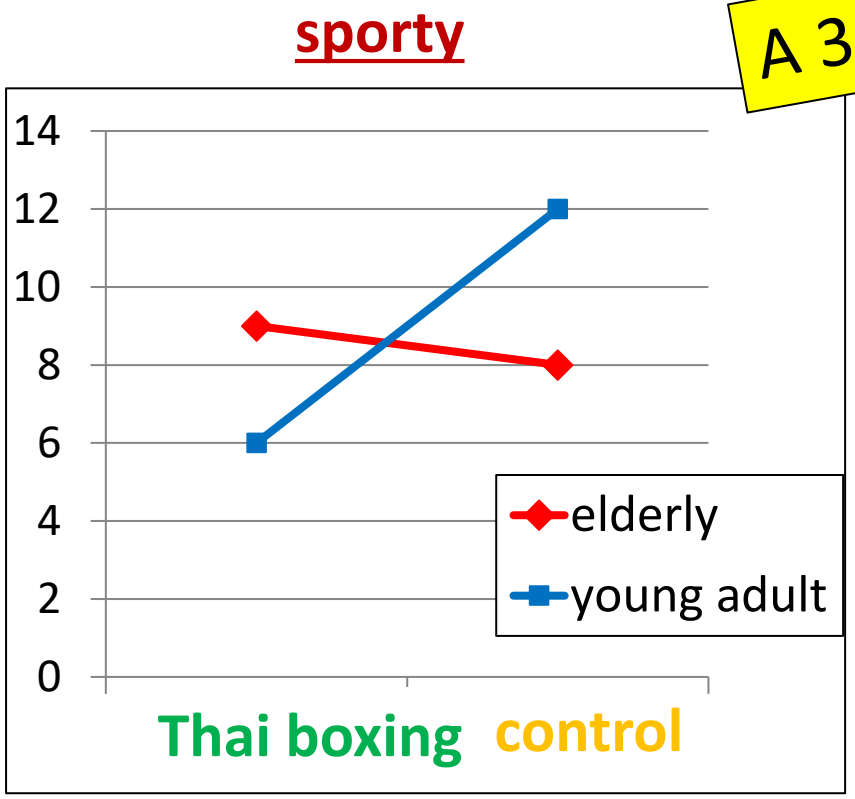
- Main effect age
- Main effect strategy
- Main effect lifestyle

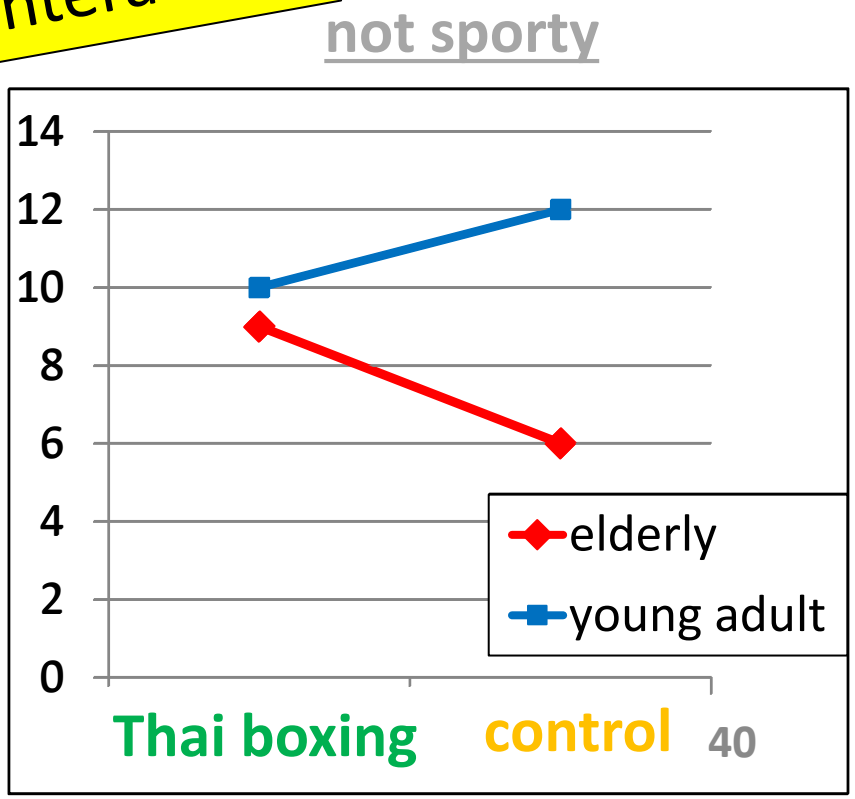
- Age x Strategy
- Age x Lifestyle
- Strategy x Lifestyle
- Age x Strategy x Lifestyle

3-way interactions

- Lifestyle x Strategy x Age
  - Interaction between ‘strategy’ and ‘age’ depends on lifestyle (sporty or not), e.g.,:

A 3-way interaction

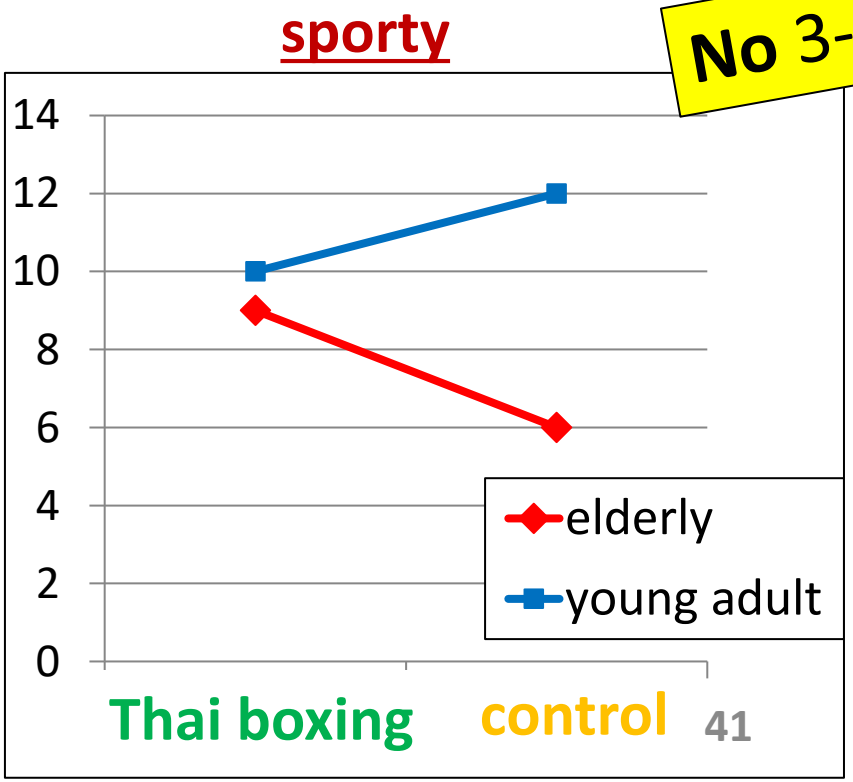


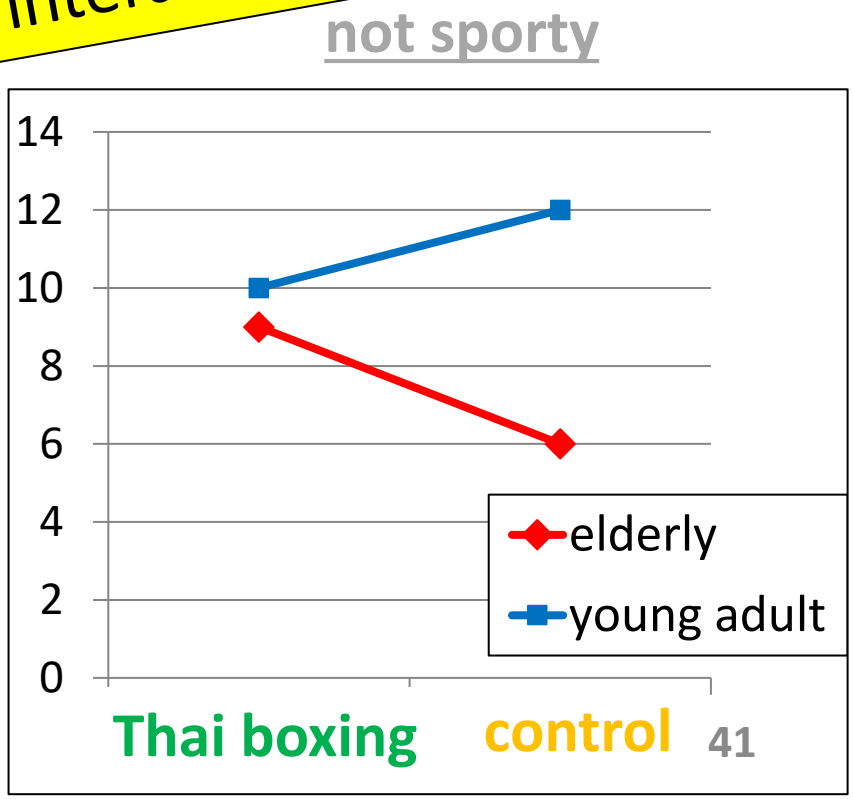


3-way interactions

- Lifestyle x Strategy x Age
  - In this case, the interaction between ‘strategy’ and ‘age’ does not depend on lifestyle (sporty or not):

No 3-way interaction





3 x 2 x 2 factorial design

Age

elderly

	Strategy	
Lifestyle	**Thai boxing**	**control**
sporty	?	?
not sporty	?	?

young adults

	Strategy	
Lifestyle	**Thai boxing**	**control**
sporty	?	?
not sporty	?	?

children

	Strategy	
Lifestyle	**Thai boxing**	**control**
sporty	?	?
not sporty	?	?



Picture source: pixabay.com

3 x 2 x 3 factorial design

Age

elderly

Strategy

Lifestyle	Thai boxing	bath	control
Sporty	?	?	?
not sporty	?	?	?

young adults

Strategy

Lifestyle	Thai boxing	bath	control
Sporty	?	?	?
not sporty	?	?	?

children

Strategy

Lifestyle	Thai boxing	bath	control
Sporty	?	?	?
not sporty	?	?	?

Today

Factorial design

- Two independent variables
  - Adding more levels
- Three independent variables
- **Factorial variations**

Obscuring factors

- Not enough variability between groups
- Too much variability within groups

Factorial variations

- Some of the factors that I included are manipulated (e.g., assigning people to conditions that differ in strategy)
- Others were not manipulated (e.g., age & lifestyle). Such variables are called *participant variables* (the levels are measured instead of manipulated)
- Both can enter a factorial design as ‘independent variable’

Factorial variations

- You can also use *within-groups* factors
  - E.g.,: first ‘taking a bath’; later ‘thai boxing’
- In case you have both independent-groups and within-groups factors: *mixed factorial design*
- Very important design to establish effectivity of a therapy
  - Thursday’s lecture

Today

Factorial design

- Two independent variables
  - Adding more levels
- Three independent variables
- Factorial variations

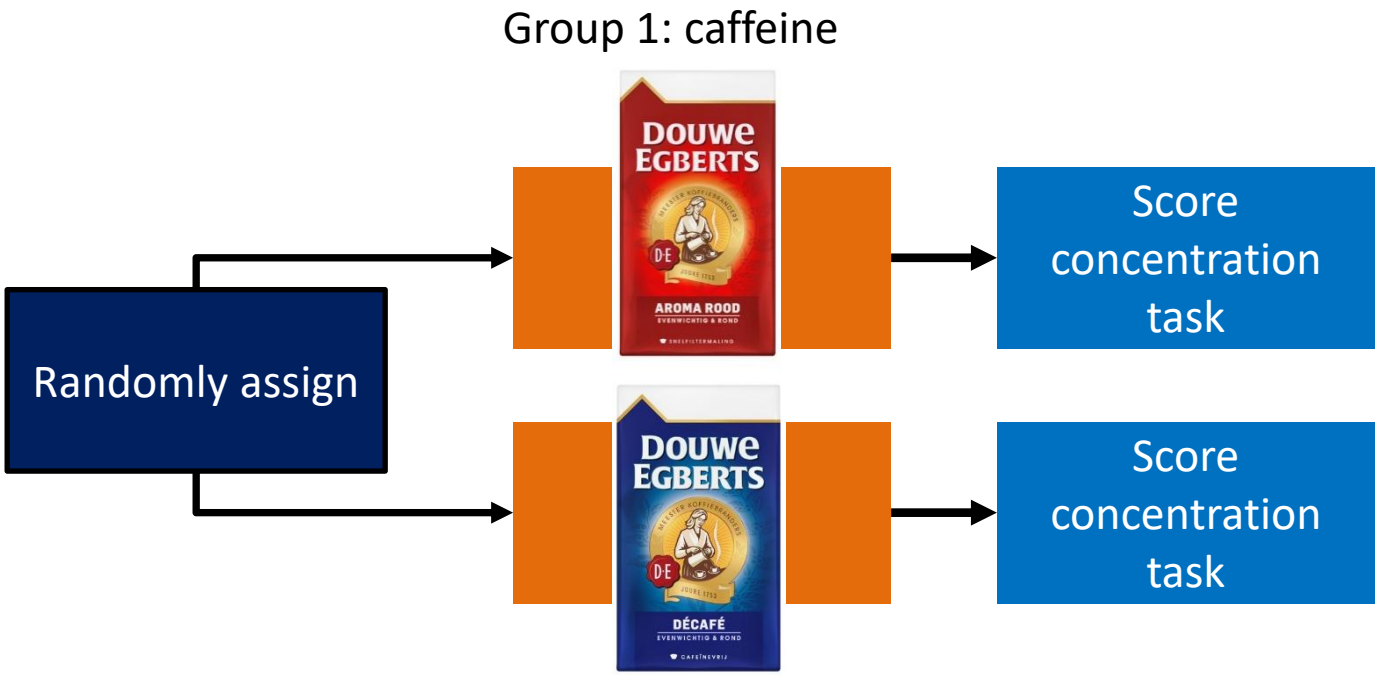
Obscuring factors

- Not enough variability between groups
- Too much variability within groups

Experiment

Influence of caffeine on concentration

Posttest-only design:



Statistical test

Is there a difference between groups?

Suppose we find 95% CI [0.14, 0.28] for the difference between groups..

Then we conclude 0 is not a believable value, so there is an effect!

There is statistical validity. But what about internal validity? (internal validity is relevant, because “caffeine influences concentration” is causal claim)

Confounding factors

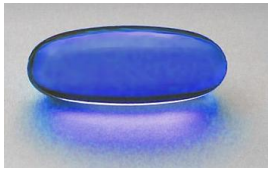
Threats to internal validity:

Is the effect on the dependent variable (e.g., concentration scores) really caused by differences in the independent variable (e.g., caffeine), or are there alternative explanations?

- For example, drinking coffee possibly confounded due to taste. So, better:



Group 1



Group 2

- Next lecture (and Chapter 11) more on confounding factors.

Photo source`pills: https://commons.wikimedia.org/wiki/File:Red_and_blue_pill.jpg

Statistical test

Is there a difference between groups?

Suppose we find 95% CI [-0.14, 0.28] for the difference between groups..

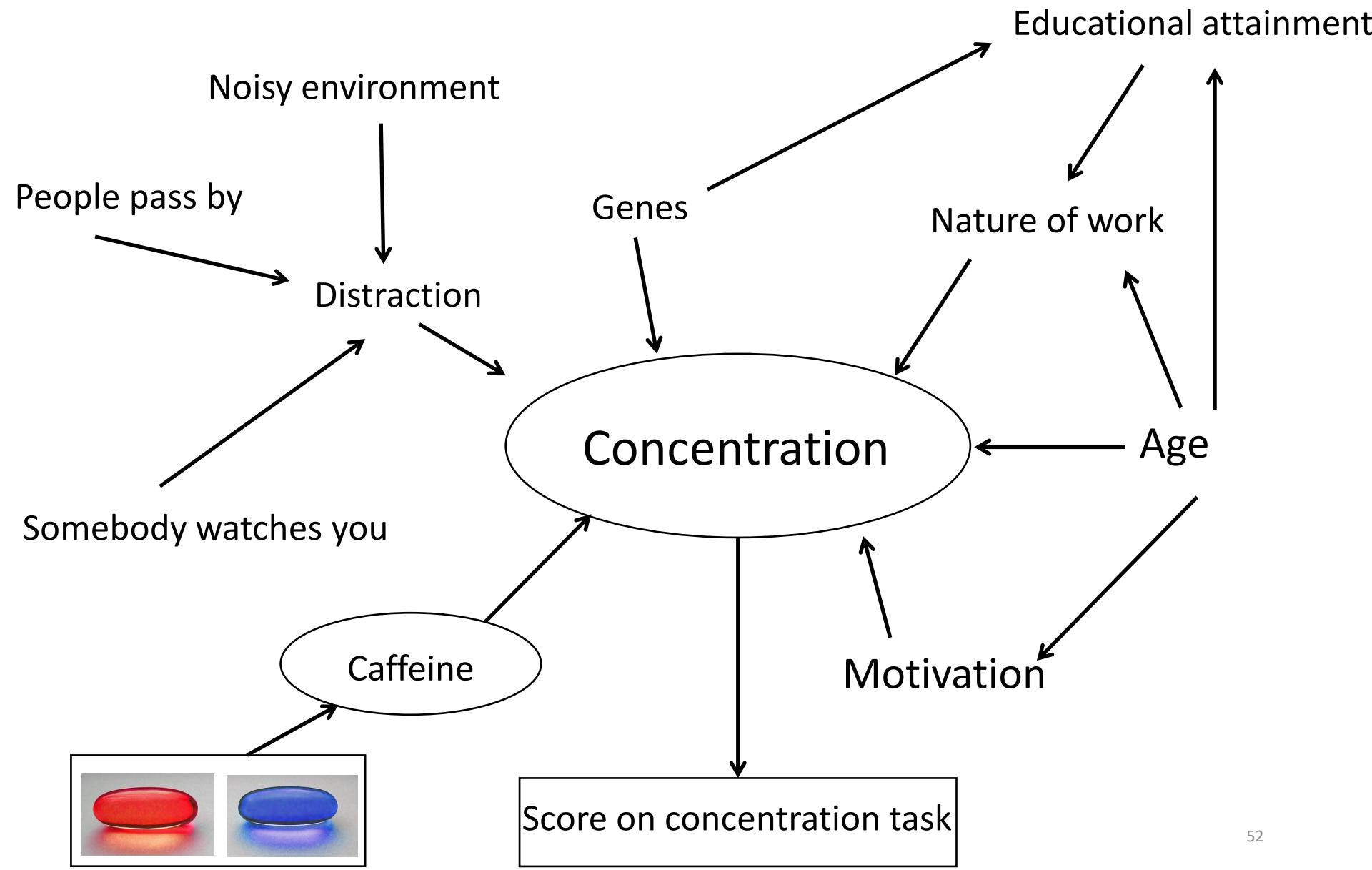
Zero is a believable value, so we cannot reject the hypothesis that there is *no difference* (difference = 0)

Such an outcome can be called a *null effect (*or ‘null result’)

Maybe there is no effect..

Or maybe there is a true effect but this study did not detect it..

Obscuring factors



Different obscuring factors

Between group issues: not enough variability between groups

- Weak manipulation
- Measurement issues

Within group issues: too much variability within groups

- Individual differences
- Situation noise
- Measurement error

Between vs within group variability

Two shakes of hot sauce:









Three shakes of hot sauce:









Picture source: pixabay.com

Between vs within group variability

Two shakes of hot sauce:









Three shakes of hot sauce:









Picture source: pixabay.com

Today

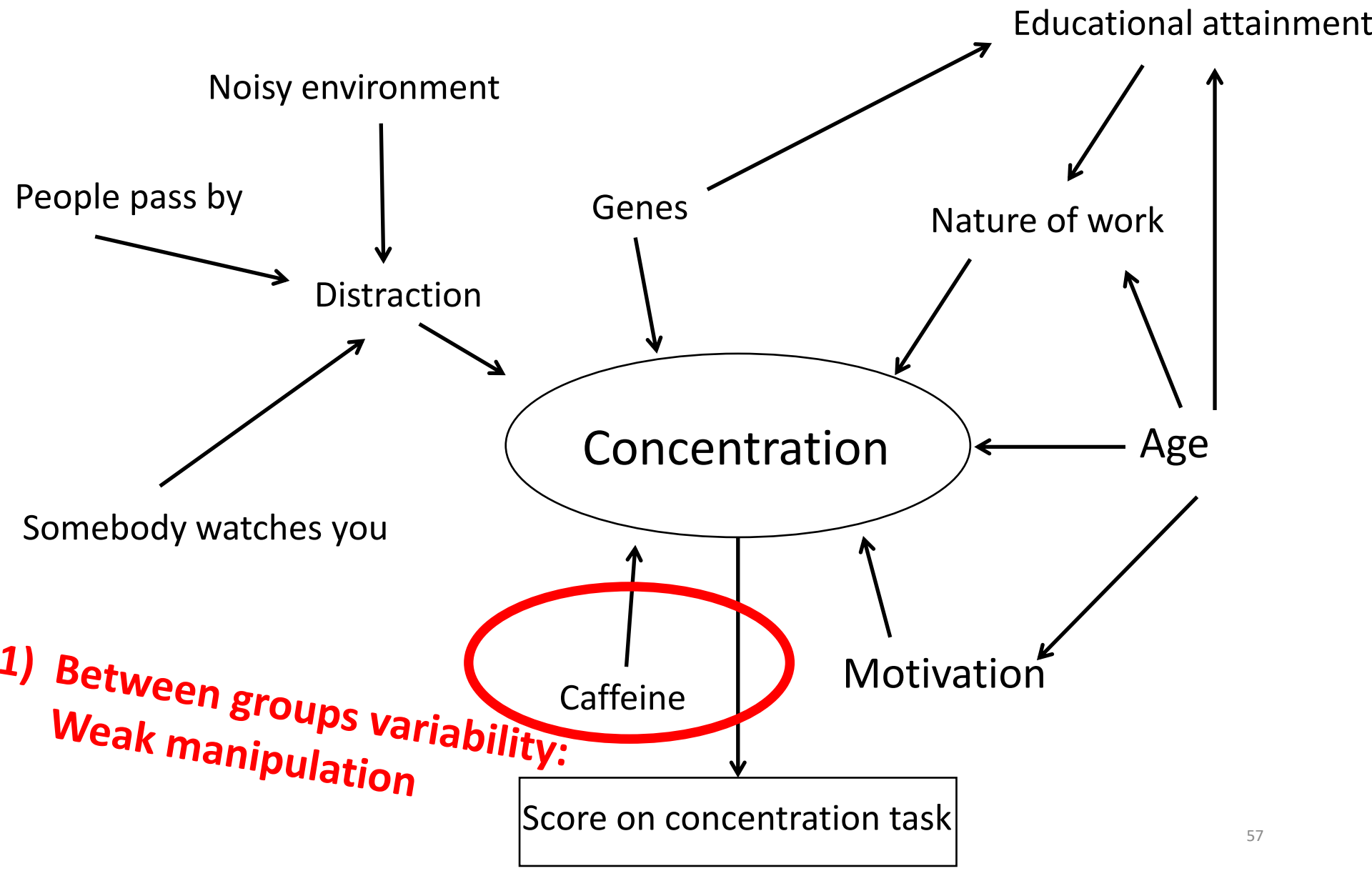
Factorial design

- Two independent variables
  - Adding more levels
- Three independent variables
- Factorial variations

Obscuring factors

- **Not enough variability between groups**
- Too much variability within groups

Obscuring factors



1) Between groups variability: Weak manipulation

Solution

Use stronger manipulations

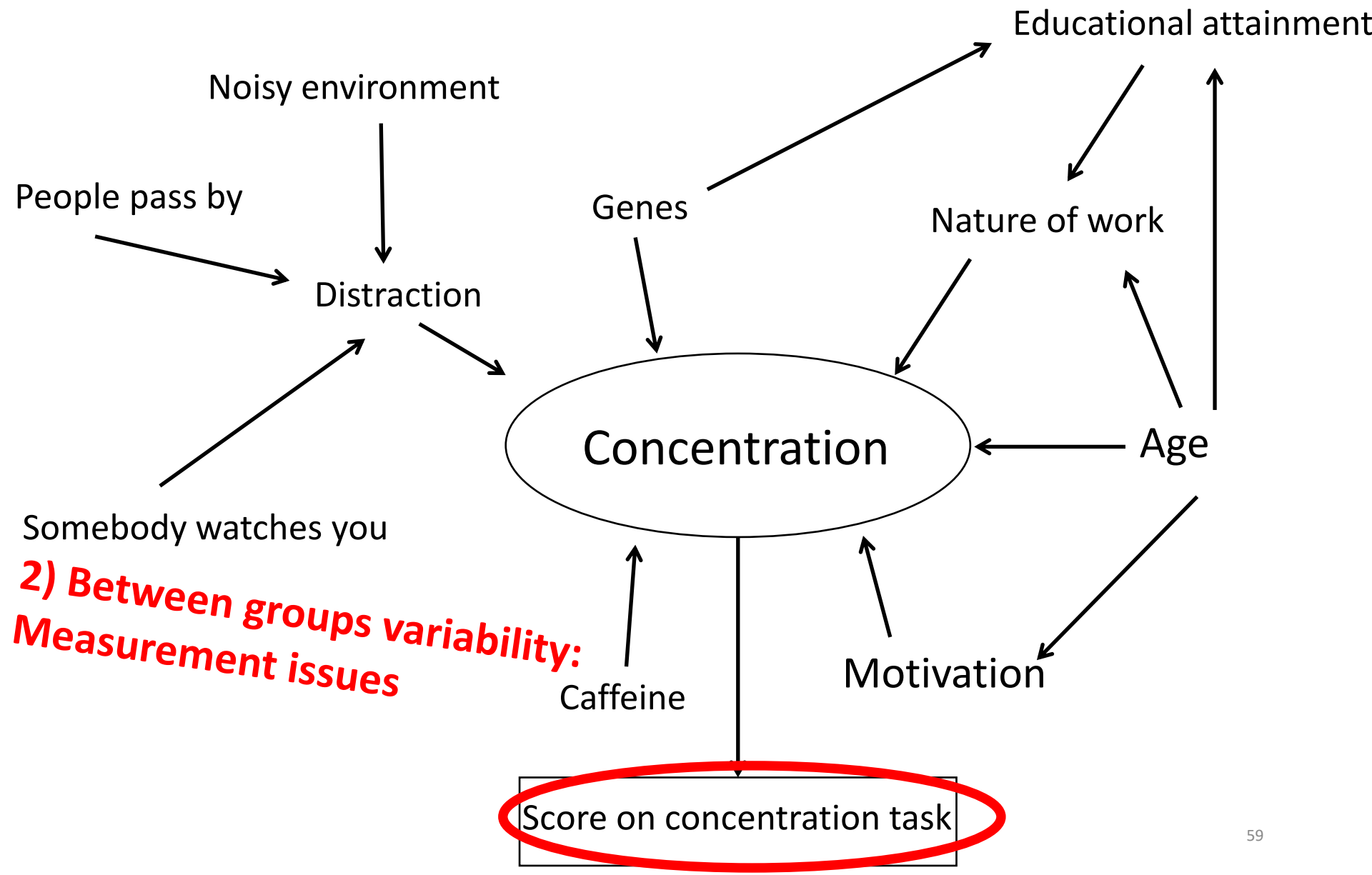




And always use a manipulation check
(see Chapter 10)

Photo source: https://www.pexels.com/photo/beverage-brown-caffeine-cappuccino-585753/

Obscuring factors



Measurement issue 1: Insensitive measures

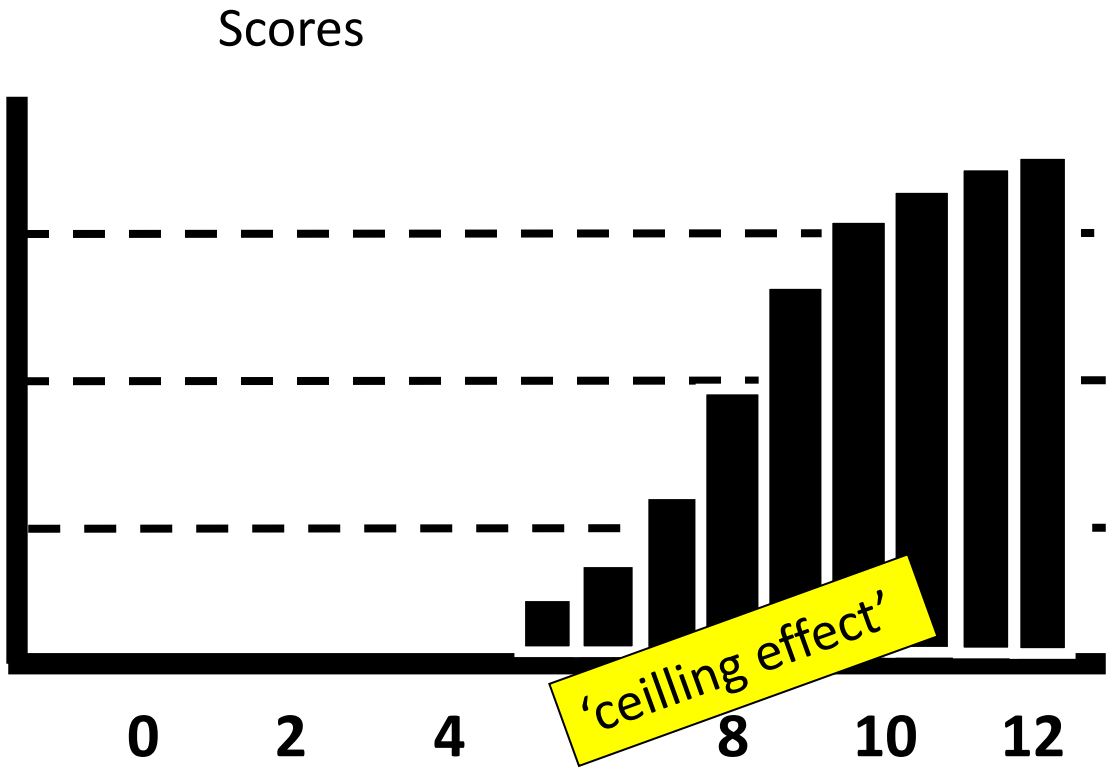
Operationalization is not sensitive enough to detect the concentration difference between the groups

→ Use a test that is more sensitive to differences in concentration

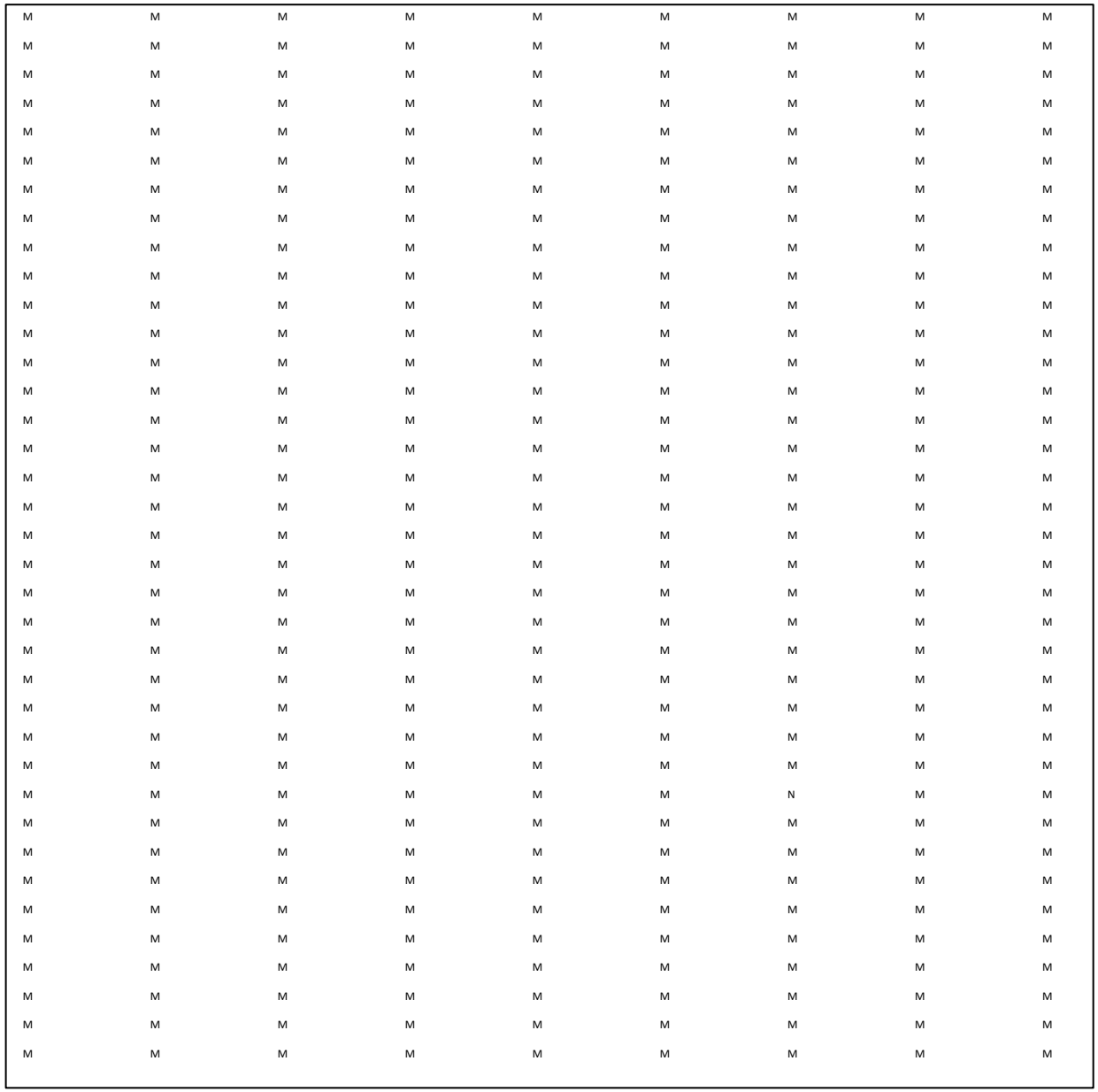
Find ‘L’

M	M	M	M	M	M	M	M	M
M	M	M	M	M	M	M	M	M
M	M	M	M	L	M	M	M	M
M	M	M	M	M	M	M	M	M
M	M	M	M	M	M	M	M	M

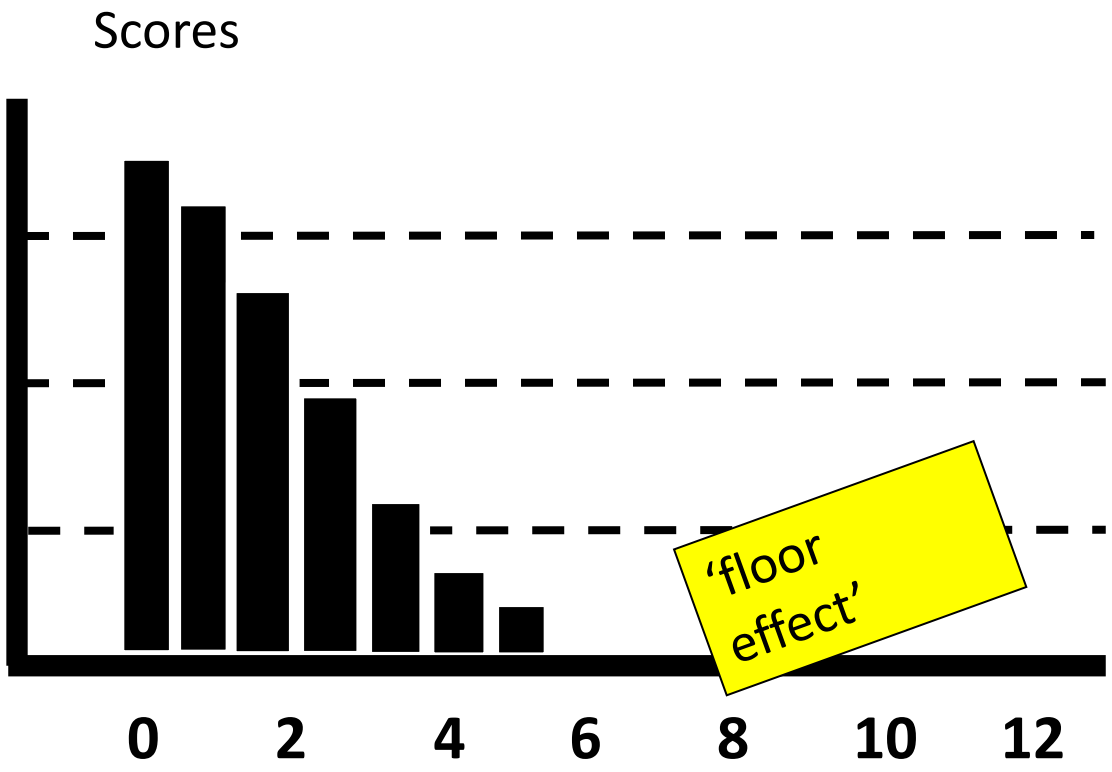
Measurement issue 2: Ceiling effects



Find ‘N’


M M M M M M M M M
M M M M M M M M M
M M M M M M M M M
M M M M M M M M M
M M M M M M M M M
M M M M M M M M M
M M M M M M M M M
M M M M M M M M M
M M M M M M M M M
M M M M M M M M M
M M M M M M M M M
M M M M M M M M M
M M M M M M M M M
M M M M M M M M M
M M M M M M M M M
M M M M M M M M M
M M M M M M M M M
M M M M M M M M M
M M M M M M M M M
M M M M M M M M M
M M M M M M M M M
M M M M M M M M M
M M M M M M M M M
M M M M M M M M M
M M M M M M N M M
M M M M M M M M M
M M M M M M M M M
M M M M M M M M M
M M M M M M M M M
M M M M M M M M M
M M M M M M M M M
M M M M M M M M M
M M M M M M M M M

Measurement issue 3: Floor effects



Today

Factorial design

- Two independent variables
  - Adding more levels
- Three independent variables
- Factorial variations

Obscuring factors

- Not enough variability between groups
- **Too much variability within groups**

Different obscuring factors

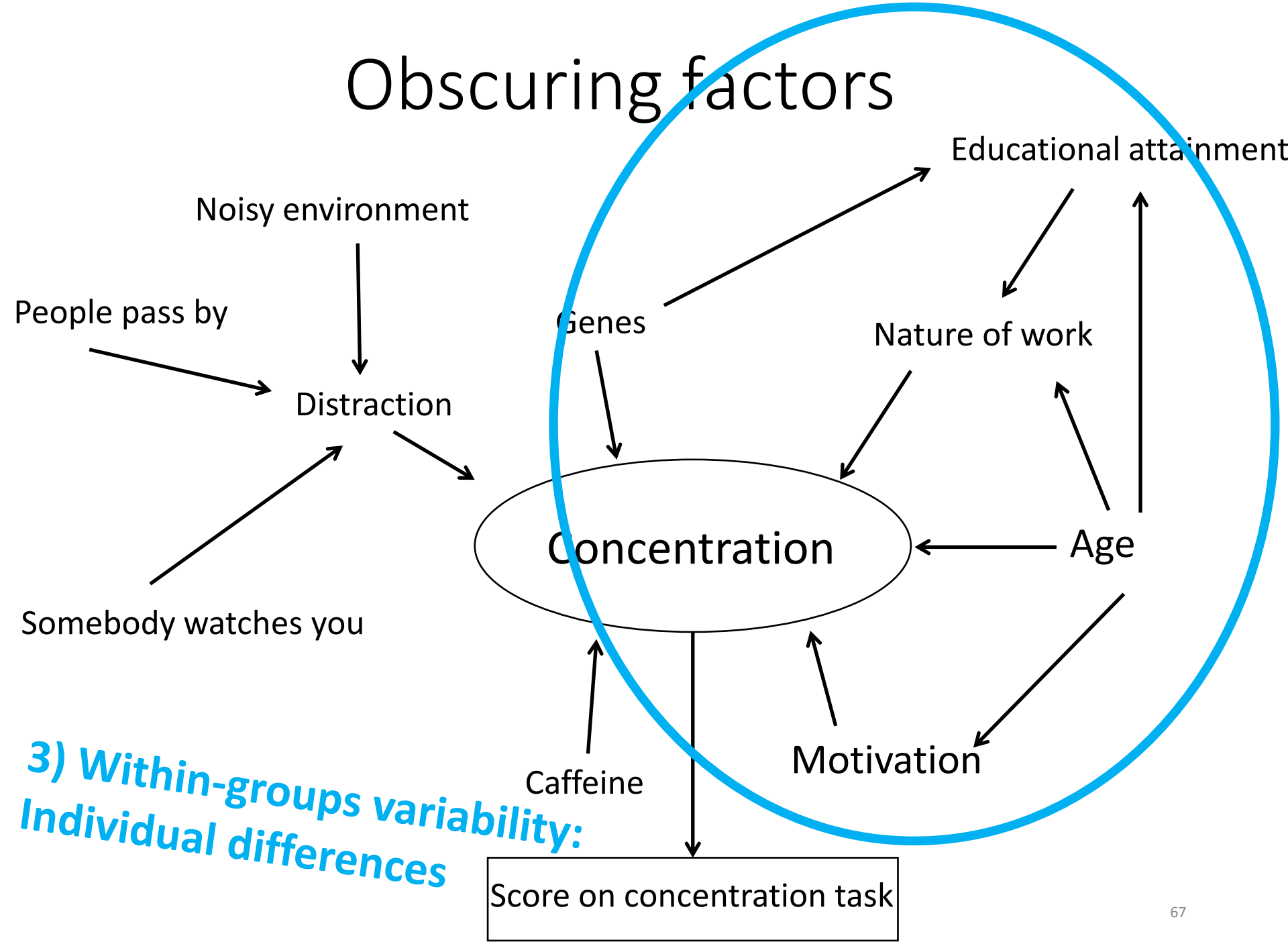
Between group issues: not enough variability between groups

- Weak manipulation
- Measurement issues

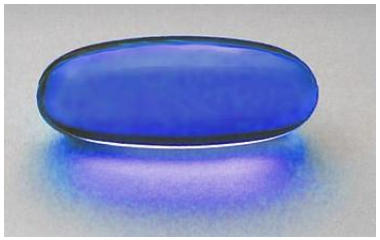
Within group issues: too much variability within groups

- Individual differences
- Situation noise
- Measurement error

Obscuring factors



3) Within-groups variability: Individual differences



Due to manipulation



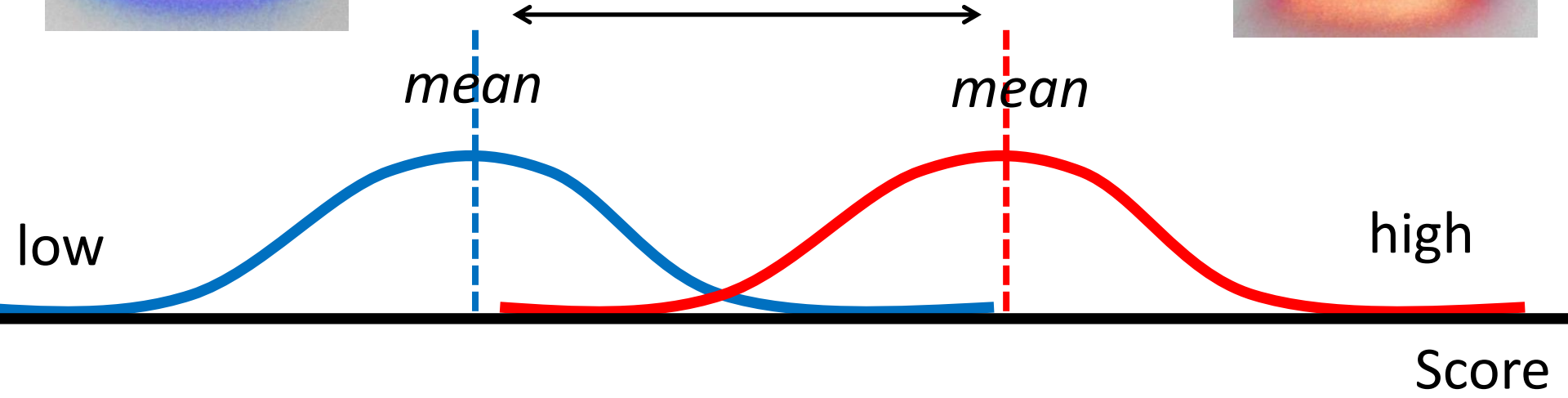
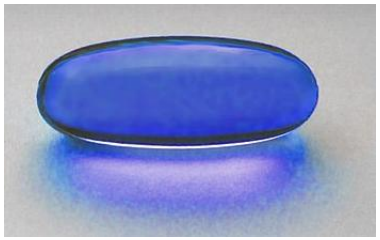


Photo source`pills: https://commons.wikimedia.org/wiki/File:Red_and_blue_pill.jpg





Due to manipulation

mean

mean

low

high

Score

Doesn't understand task

Not motivated

Has a job that involves concentration

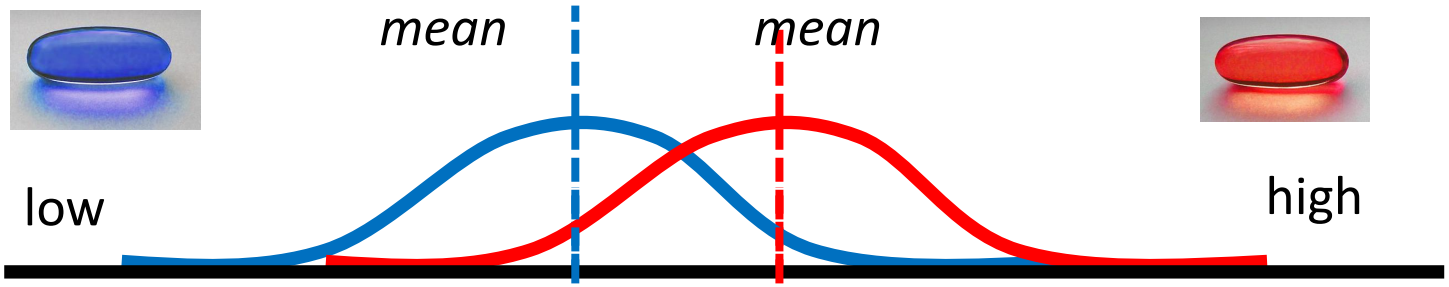
Favorable genes

Variation within groups due to individual differences

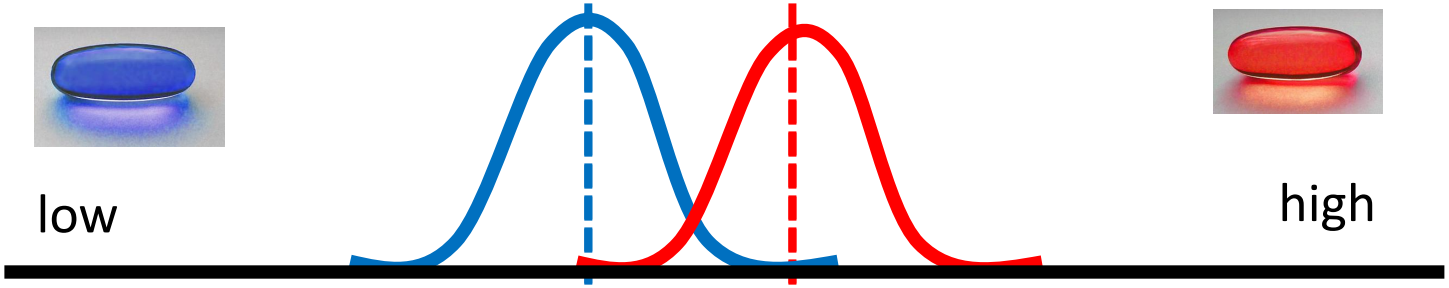
Photo source`pills: https://commons.wikimedia.org/wiki/File:Red_and_blue_pill.jpg

2) Within-groups issue: individual differences

More individual differences



Less individual differences



Less individual differences: Better detectable effect!

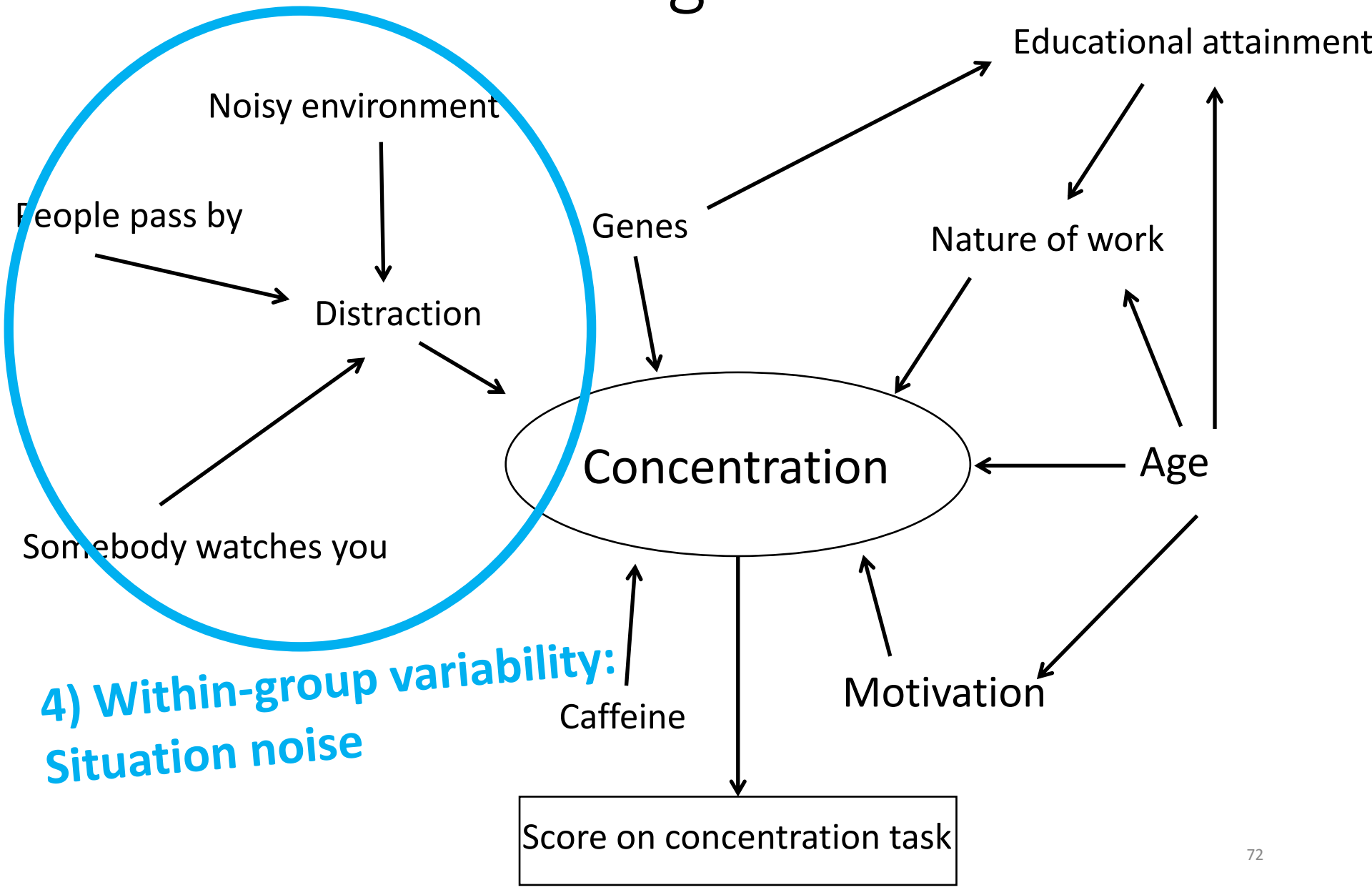
Photo source`pills: https://commons.wikimedia.org/wiki/File:Red_and_blue_pill.jpg

Solutions

- Use a within-groups design (Chapter 10)
  - As this design controls for individual differences
- Control important individual differences variables
  - E.g., control age and usual coffee intake, or do not include people with large concentration issues
- Add more participants
  - As it will decrease the standard error of the mean:

$$\frac{s}{\sqrt{n}}$$

Obscuring factors



4) Within-group variability: Situation noise

Solution

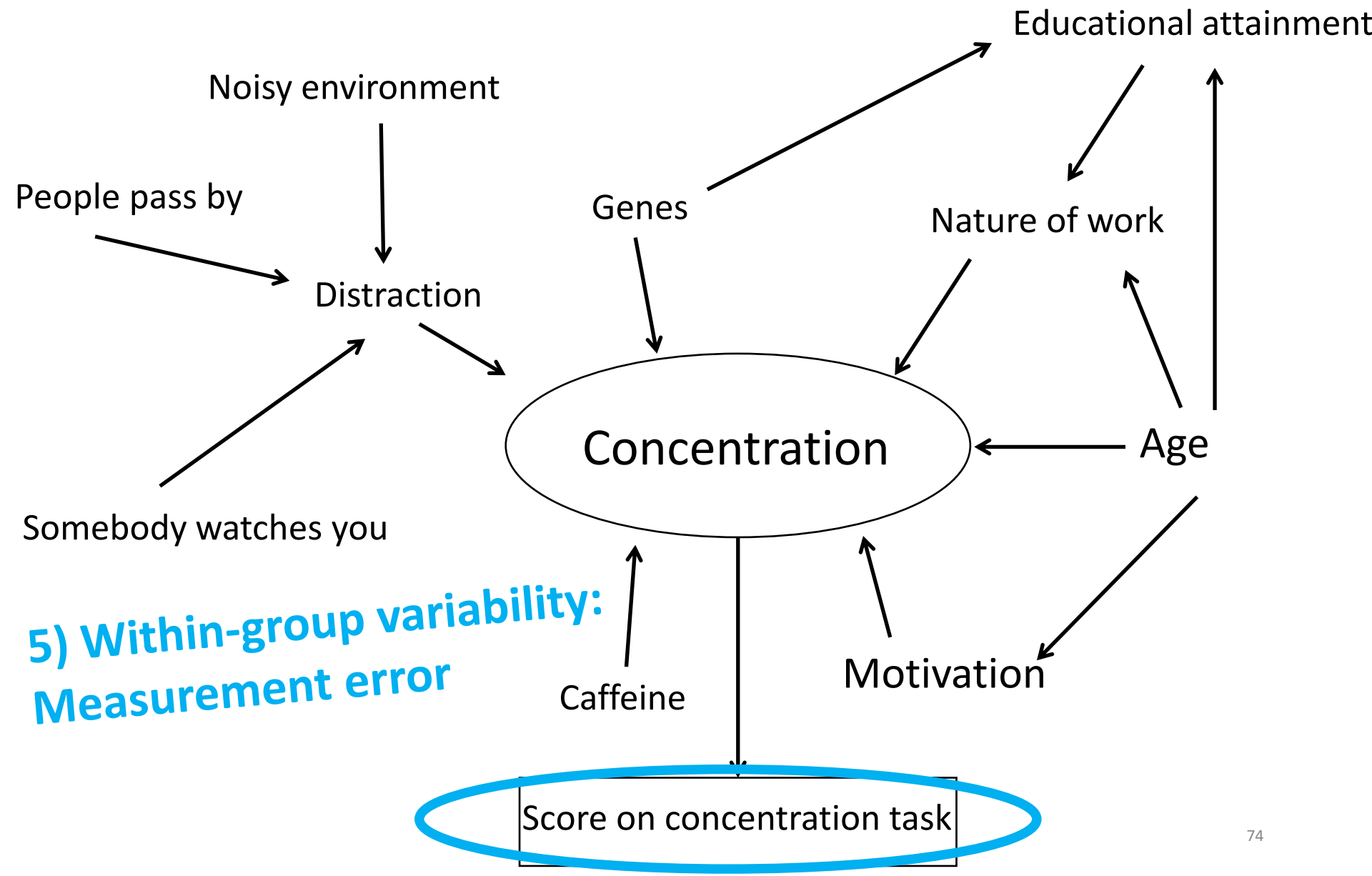
- Control what you can control
  - e.g., don't use rooms in loud environments
  - e.g., don't stress you participants
  - etc

- For those things you can't control, at least make sure it's random
  - e.g., if you have two testing rooms, assign every participant at random to one of the rooms

  →i.e., not one condition in the one room, and another condition in the other room

  Otherwise such factors can become confounders!

Obscuring factors



Reliable scores
(reasonably consistent scores)

Reliable?

		Concentration	M1	M2	M3
	Jordan	5	5	4	6
	Lady Gaga	4	3	5	4
	Peter	3	2	2	4
	Dopey	1	1	2	0
	Serena	4	4	5	3

2) Within-groups issue: measurement error

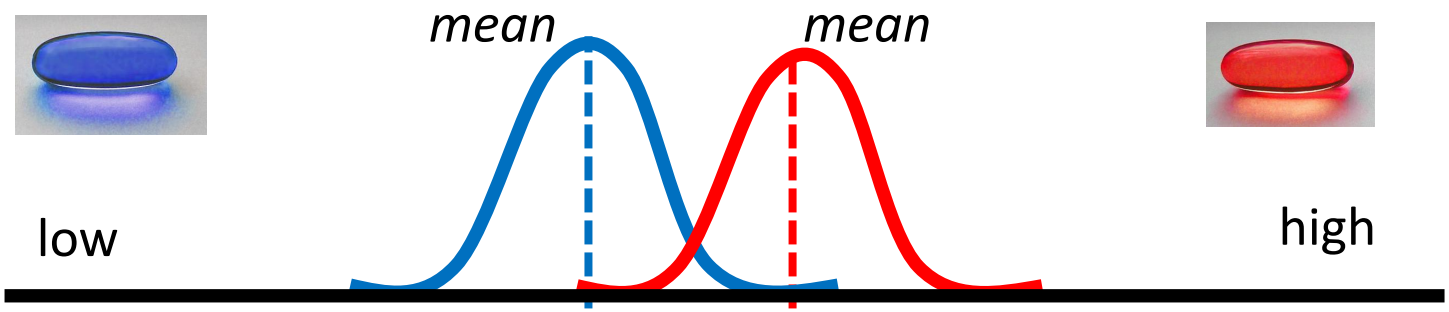
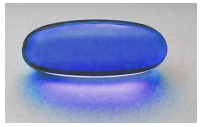


Photo source`pills: https://commons.wikimedia.org/wiki/File:Red_and_blue_pill.jpg

2) Within-groups issue: measurement error

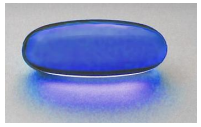


Perfectly reliable concentration task

Photo source`pills: https://commons.wikimedia.org/wiki/File:Red_and_blue_pill.jpg

2) Within-groups issue: measurement error

More variation due to measurement error!

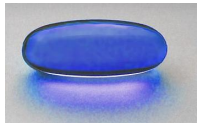


Reasonably reliable concentration task

Photo source`pills: https://commons.wikimedia.org/wiki/File:Red_and_blue_pill.jpg

2) Within-groups issue: measurement error

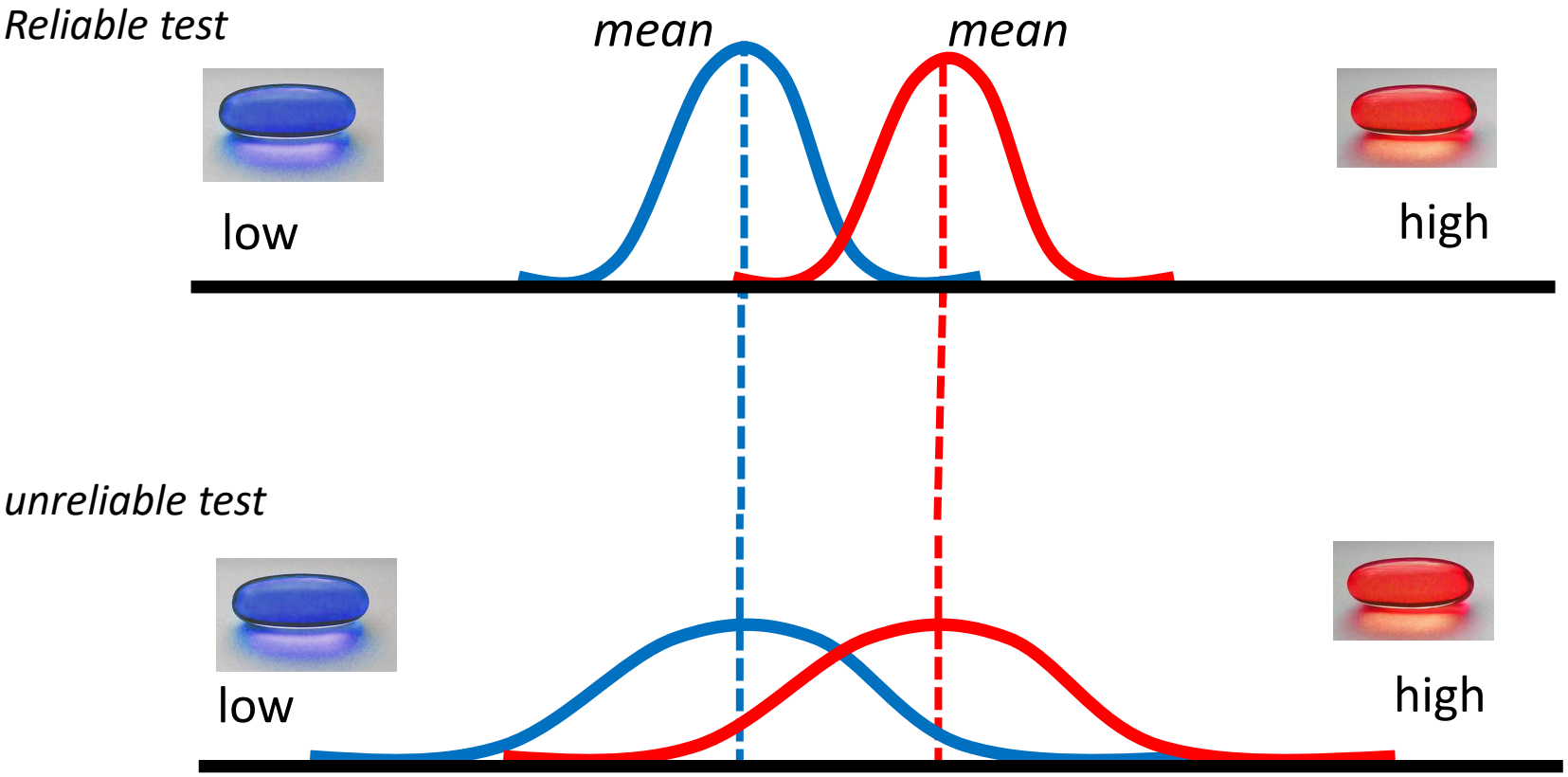
Even more variation due to measurement error!



Very **un**reliable concentration task

Photo source`pills: https://commons.wikimedia.org/wiki/File:Red_and_blue_pill.jpg

2) Within-groups issue: measurement error



More measurement error: Less clear effect!

Photo source`pills: https://commons.wikimedia.org/wiki/File:Red_and_blue_pill.jpg

Solution

- Use reliable instruments (good internal, interrater and test-retest reliability)

- Use more items and/or more participants to reduce the effect of measurement error

Summary:

- When the difference that an independent variable makes on the dependent variable depends on another independent variable, there is an interaction between the independent variables.
- Factorial design powerful tool to test for interaction effects
- In setting up an experiment, be aware of potential *obscuring factors*: These are factors that make it difficult to detect an effect when there is one (alternative explanations for a null result)
- Some factors obscure an effect because there is too little difference *between* groups, and other factors obscure the effect because there is too much variability *within* the groups

Example exam questions

A researcher uses a 2x2x2 factorial design. How many independent variables does this design have?

A 2 independent variables
B 3 independent variables
C 8 independent variables

Consider the following experimental design with 2 between-subject factors: The first factor has 4 levels, the second factor has 2 levels. How many conditions does this design have?

A 4 conditions
B 6 conditions
C 8 conditions

Answers Example exam questions

A researcher uses a 2x2x2 factorial design. How many independent variables does this design have?

A 2 independent variables
B 3 independent variables
C 8 independent variables

count the number of numbers: 2x2x2 means three independent variables with each two levels

Consider the following experimental design with 2 between-subject factors: The first factor has 4 levels, the second factor has 2 levels. How many conditions does this design have?

A 4 conditions
B 6 conditions
C 8 conditions

Multiply: 2 x 4 = 8

Example exam question

Amelia is interested in the effect of certain morning routines, such as drinking coffee and eating breakfast, on energy levels later in the day. Consider the following results of an Independent-groups design with two independent variables (eating breakfast or not, drinking coffee or not) and one dependent variable (energy level). Which of the following conclusions is correct? You can ignore sampling variability and only consider the point estimates in the table.

		breakfast	
coffee		**yes**	**no**
	yes	12.3	11.1
	no	12	11.4

A There is a main effect of coffee and an interaction effect
B There is a main effect of breakfast and an interaction effect
C There are main effects but no interaction effect

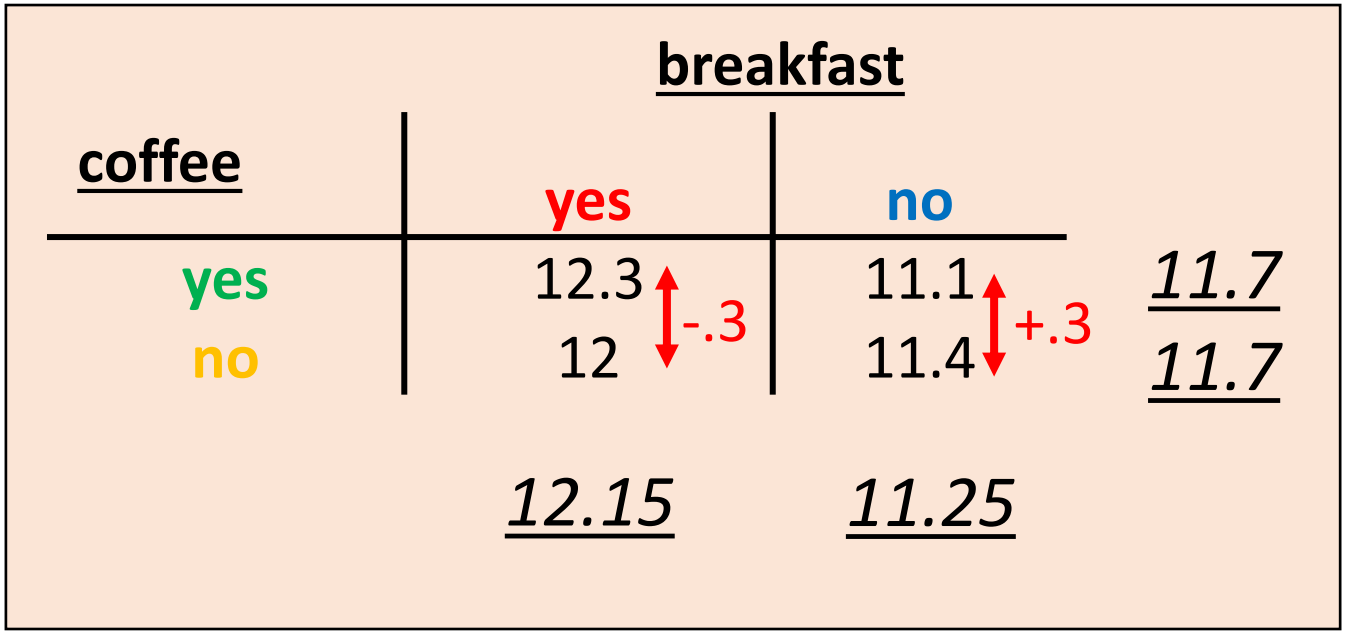
Example exam question

Amelia is interested in the effect of certain morning routines, such as drinking coffee and eating breakfast, on energy levels later in the day. Consider the following results of an Independent-groups design with two independent variables (eating breakfast or not, drinking coffee or not) and one dependent variable (energy level). Which of the following conclusions is correct? You can ignore sampling variability and only consider the point estimates in the table.

		breakfast
coffee	**yes**	**no**
yes	12.3	11.1
no	12	11.4

A There is a main effect of coffee and an interaction effect
B There is a main effect of breakfast and an interaction effect
C There are main effects but no interaction effect

Answer Example exam question



11.7 = 11.7 so there is no main effect of coffee
12.15 ≠ 11.25 so there is a main effect of breakfast
-0.3 ≠ 0.3 so there is a main effect of breakfast

A There is a main effect of coffee and an interaction effect
B There is a main effect of breakfast and an interaction effect
C There are main effects but no interaction effect