

Mediation

- explains how or why an intervention works
- mediator explains all or part of the treatment's impact on an intended outcome
- is an intermediate outcome that is measured or observed after the onset of the intervention. E.g. fidelity of application, how many questions were asked ?

Moderation

- explains who the intervention benefits or what conditions must exist for the intervention to be effective.
- a factor that reflects who is most affected by the treatment
- a factor that exists prior to the introduction of an intervention Eg. student characteristics, such as special education status, gender, ...

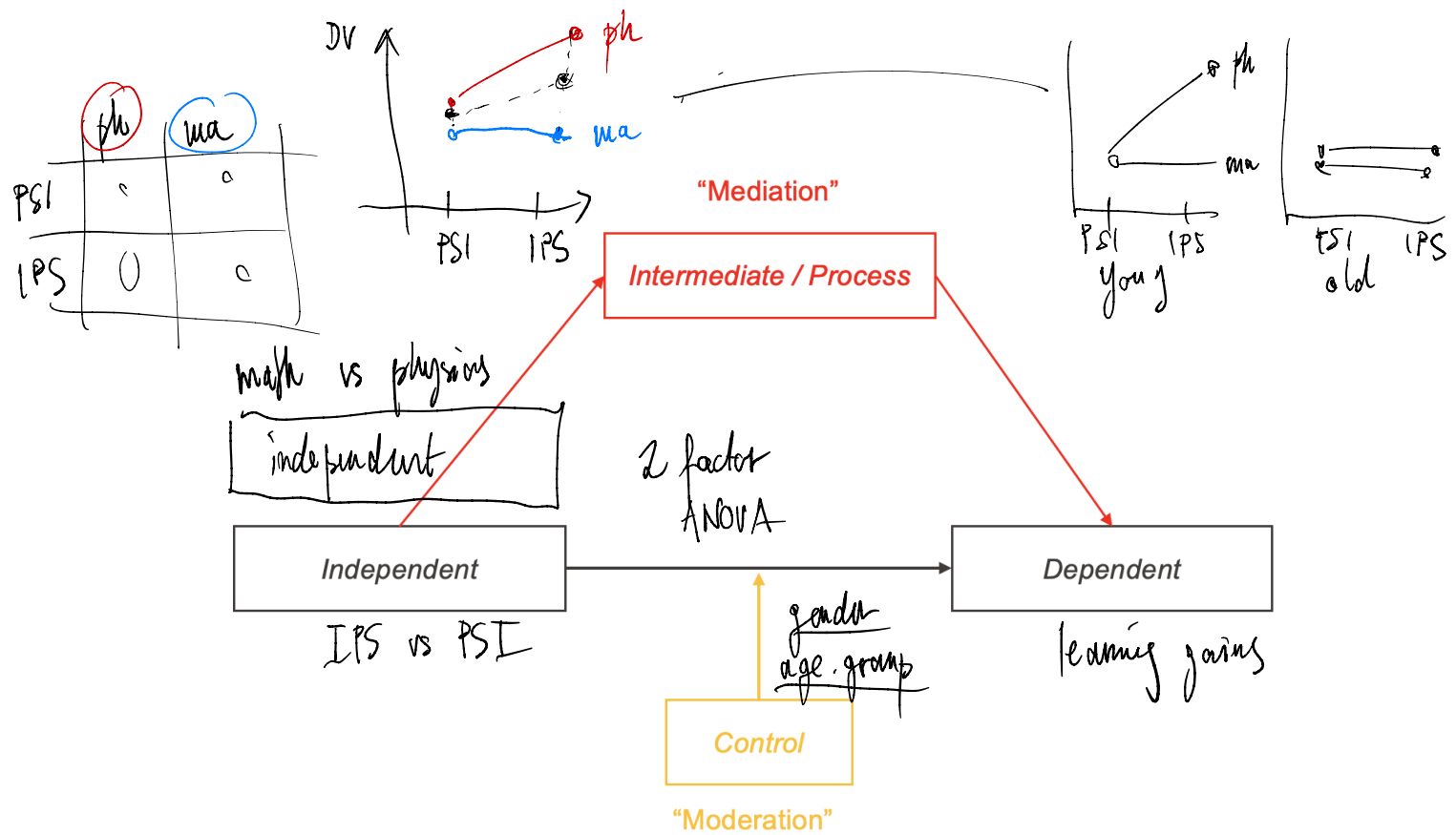
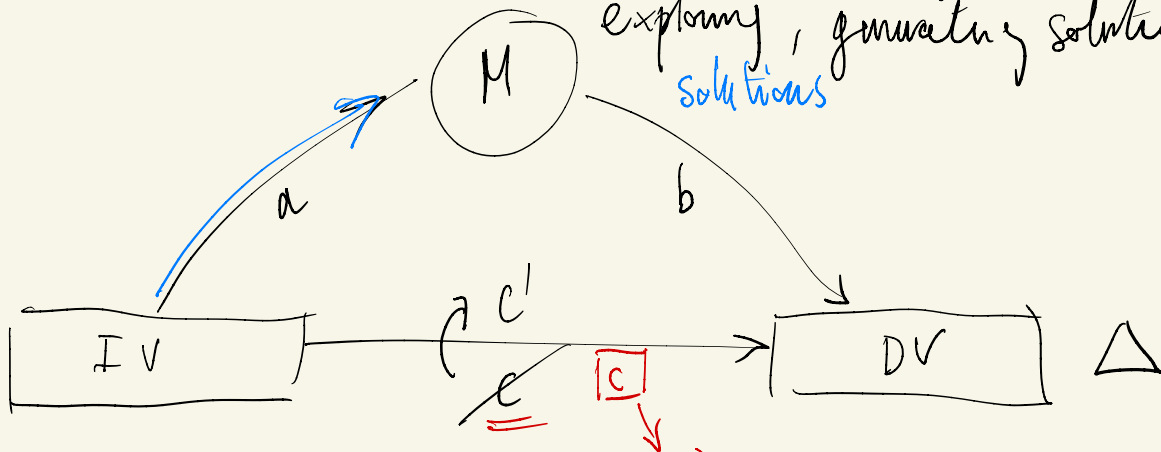


Figure 1: Overview

process variable building hypothesis
 discuss more, argue
 exploring, generating solutions



$$c = \lim (dv \sim iv)$$

$$a = \lim (\text{solutions} \sim iv)$$

$$c' = \lim (dv \sim \underset{\substack{\uparrow \\ c'}}{iv} + \underset{\substack{\uparrow \\ b}}{\text{solutions}})$$

$$c > c'$$

direct

Dataset

This dataset was generated to illustrate basic statistical techniques like ANOVA and regression as well lightly more advanced techniques like mediation and moderation. However, we tried as much as possible to implement variations compatible with insights found in the literature about Productive Failure:

Sinha, T., & Kapur, M. (2021). When Problem Solving Followed by Instruction Works: Evidence for Productive Failure. *Review of Educational Research*, 91(5), 761–798. <https://doi.org/10.3102/00346543211019105>

Participants

The sample consisted of N=200 participants.

Independent variable

Order of instruction The independent variable has two modalities (also called conditions):

- I-PS : instruction followed by problem-solving
- PS-I : problem-solving followed by instruction

Participants were *randomly* assigned to one of the experimental conditions.

Dependent variable

Learning gain. Participants completed a 10 question *pre-test* before starting the experiment. The pre-test was a series of questions about their understanding of the sun-earth relative positions. After the experiment, participants completed a 10 question *post-test* with similar questions as the pre-test. The learning gain was computed as :

$$learning.gain = post.test - pre.test$$

another possibility would be the relative learning gain

$$rel.gain = \frac{post.test - pre.test}{max - pre.test}$$

Control variables

Age group. Participants were recruited among highschool students who are interested in following studies at EPFL (kids), students doing their bachelor as well as alumni who are active professionally (professionals).

Young learners (e.g., second to fifth graders) may have insufficient prior knowledge about cognitive and metacognitive learning strategies to generate multiple solutions during initial problem solving

Gender. Experimenters also asked for the gender of the participants, either Male (M) or Female (M).

Self-regulation skills. Participants also filled in a questionnaire about their self-regulation skills by using the Learning Companion (<https://companion.epfl.ch>)

Intermediate / Process variables

Solutions. The simulation system logged every simulation run and counted how often students used the simulation to generate a potential solution.

Mediation

- explains how or why an intervention works
- mediator explains all or part of the treatment's impact on an intended outcome
- is an intermediate outcome that is measured or observed **after** the onset of the intervention. E.g. fidelity of application, how many questions were asked ?
- there is a plausible causality relation between the experimental treatment and the mediating variable.

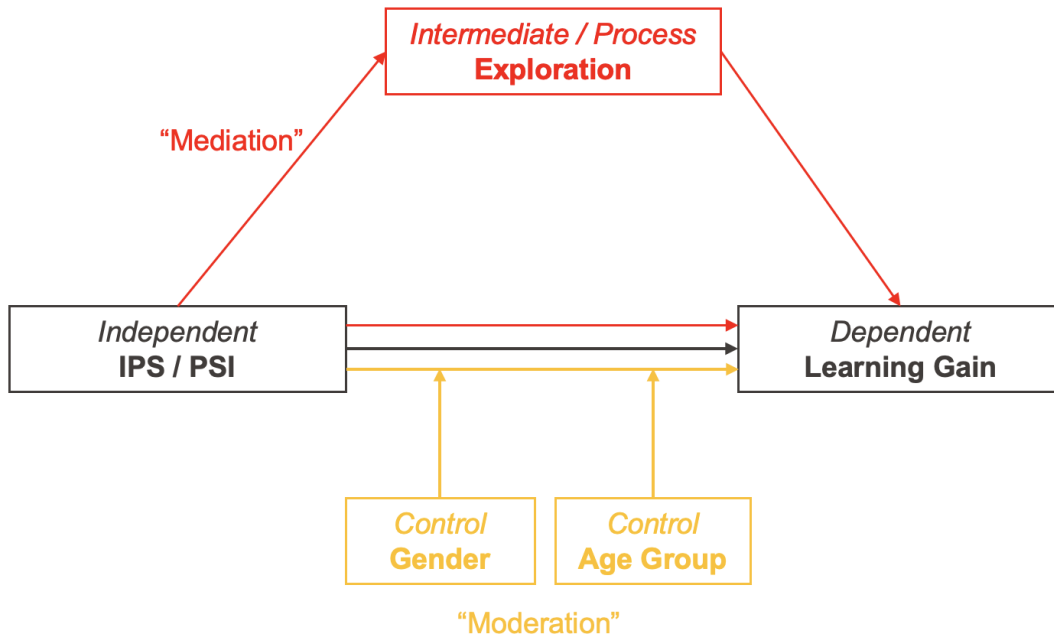


Figure 5: Overview

The “old” way

Baron and Kenny’s (1986) steps for mediation analysis See <https://davidakenny.net/cm/mediate.htm>

Step 1: Show that the causal variable is correlated with the outcome. Use Y as the criterion variable in a regression equation and X as a predictor (estimate and test path c in the above figure). This step establishes that there is an effect that may be mediated.

Step 2: Show that the causal variable is correlated with the mediator. Use M as the criterion variable in the regression equation and X as a predictor (estimate and test path a). This step essentially involves treating the mediator as if it were an outcome variable.

Step 3: Show that the mediator affects the outcome variable. Use Y as the criterion variable in a regression equation and X and M as predictors (estimate and test path b). It is not sufficient just to correlate the mediator with the outcome because the mediator and the outcome may be correlated because they are both caused by the causal variable X. Thus, the causal variable must be controlled in establishing the effect of the mediator on the outcome.

Step 4: To establish that M completely mediates the X-Y relationship, the effect of X on Y controlling for M (path c’) should be zero (see discussion below on significance testing). The effects in both Steps 3 and 4 are estimated in the same equation.

$c \text{ model} = \text{lm}(\text{learning} \sim \boxed{\text{condition}})$
 $a \text{ model} = \text{lm}(\text{solutions} \sim \text{condition})$
 $c' \text{ model} = \text{lm}(\text{learning} \sim \boxed{\text{condition}} + \boxed{\text{solutions}})$

mediator

if mediation
 then
 condition is not sig
 or is much weaker

"Mediation"

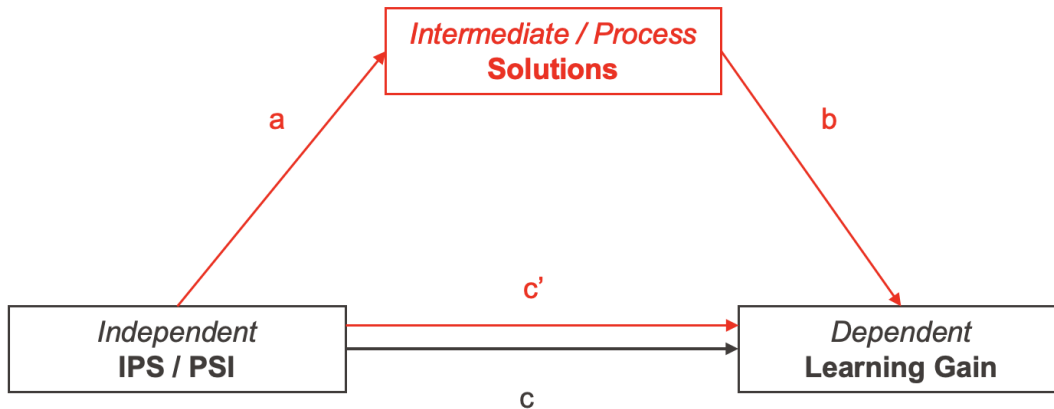
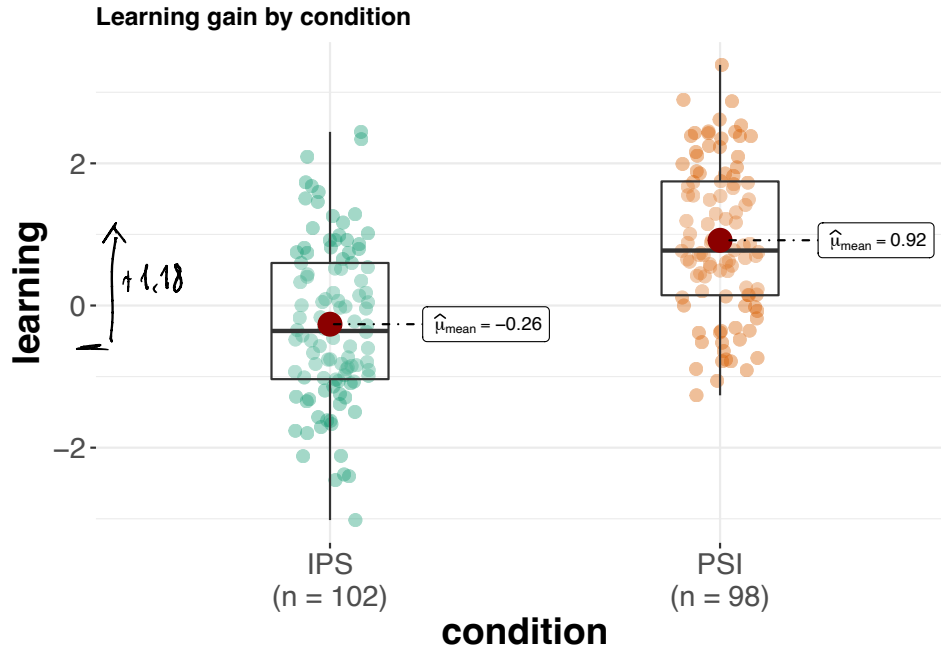
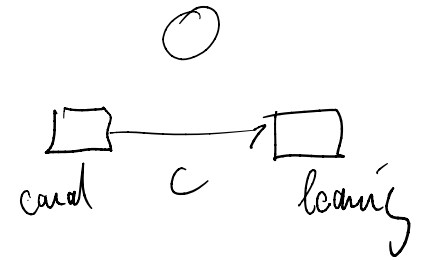


Figure 7: Mediation

Condition affects the learning gain (c path)



Step 1

First we make sure the experimental treatment affects the dependent variable.

```
c.path <- lm(learning ~ condition, data=df)
summary(c.path)
```

```
##
## Call:
## lm(formula = learning ~ condition, data = df)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.7555 -0.7754 -0.1243  0.8605  2.7054
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -0.2626     0.1092  -2.405   0.0171 *
## conditionPSI    1.1806     0.1560   7.567 1.41e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.103 on 198 degrees of freedom
## Multiple R-squared:  0.2243, Adjusted R-squared:  0.2204
## F-statistic: 57.26 on 1 and 198 DF, p-value: 1.412e-12
```

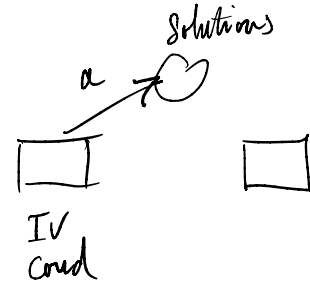
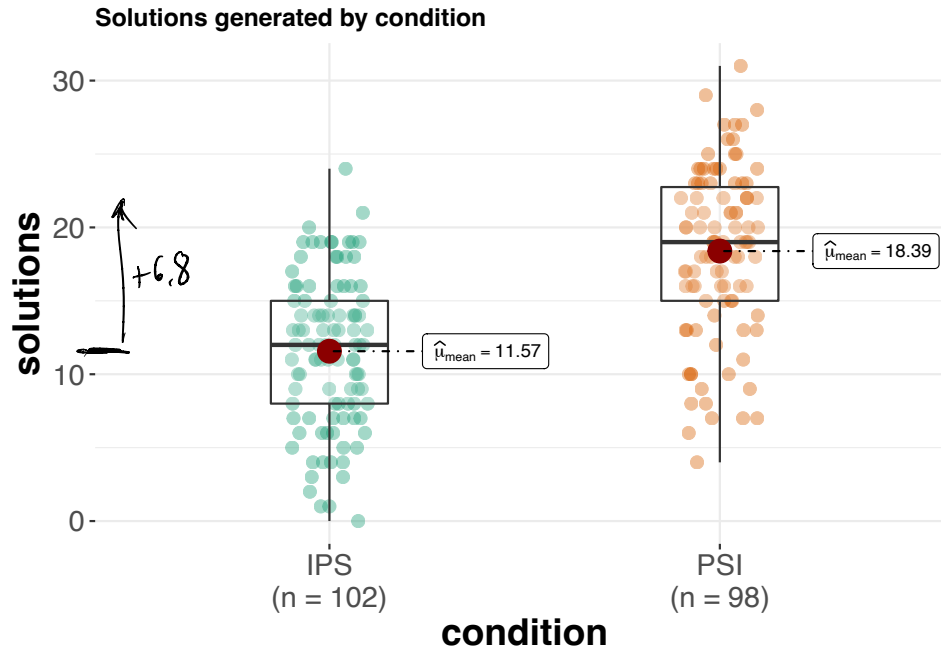
$$y = ax + b + \epsilon$$



$$\beta_{\text{cond}} = 1.18$$

% of variance explained

Condition affects the number of solutions generated (a path)



Step 2

This model checks whether the experimental treatment affects the mediating variable

```
a.path <- lm(solutions ~ condition, data=df)
summary(a.path)
```

```
##
## Call:
## lm(formula = solutions ~ condition, data = df)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -14.3878  -3.5686   0.5218   3.6122  12.6122
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    11.5686     0.5344   21.647 < 2e-16 ***
## conditionPSI     6.8191     0.7635    8.932 2.9e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.397 on 198 degrees of freedom
## Multiple R-squared:  0.2872, Adjusted R-squared:  0.2836
## F-statistic: 79.78 on 1 and 198 DF, p-value: 2.903e-16
```

$$\beta_{cond} = 6.81$$

Step 3 & 4

Finally we check whether:

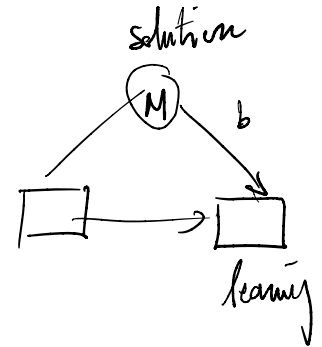
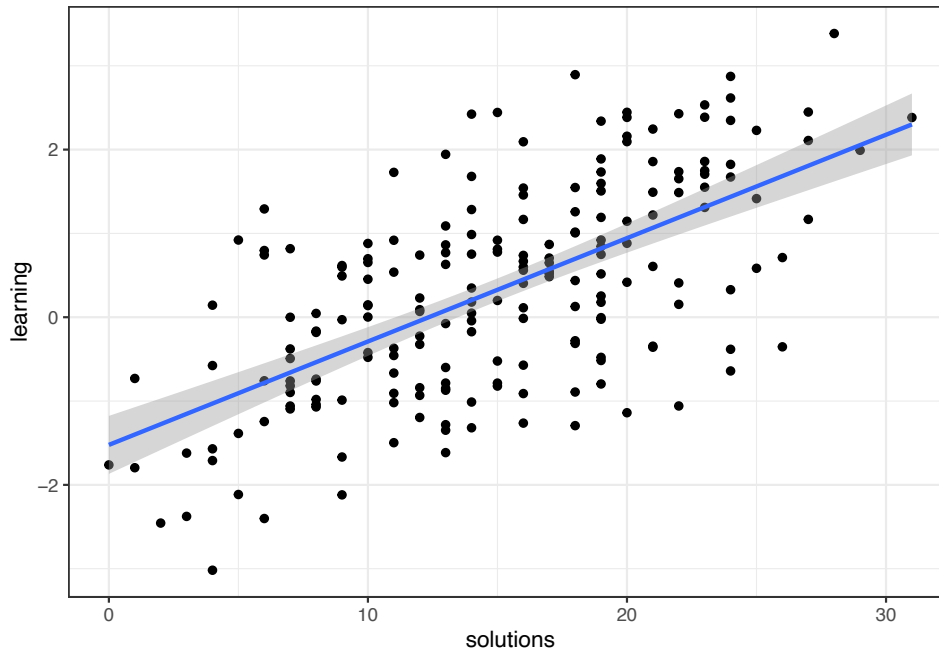
- the mediating variable affects the dependent variable and
- that the effect of the independent variable **decreases** (partial mediation) or even **disappears** (full mediation).

In our case:

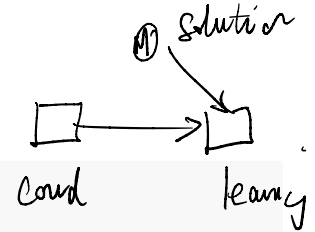
- the Estimate for the mediator (solutions) is statistically significant and
- the Estimate for condition PSI went down from 1.1806 to 0.4764, but is still significant => we have a partial mediation

Solutions and Learning gain are correlated (b path)

[1] 0.6295641



mediator



```
c.dash.path <- lm(learning ~ condition + solutions, data=df)
summary(c.dash.path)
```

```
##
## Call:
## lm(formula = learning ~ condition + solutions, data = df)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.34975 -0.64637 -0.00054  0.65168  2.35102
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -1.45730    0.17336   -8.406 8.41e-15 ***
## conditionPSI  0.47639    0.15987    2.980 0.00325 **
## solutions     0.10327    0.01256    8.219 2.70e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9542 on 197 degrees of freedom
## Multiple R-squared:  0.4224, Adjusted R-squared:  0.4165
## F-statistic: 72.03 on 2 and 197 DF, p-value: < 2.2e-16
```

$C = 1.18$
 $C' = 0.47$

41%

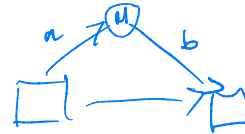
partial mediation
 condition still signif

Mediation - Sobel test.

Sobel has developed a method to test whether this mediation effect is significant. See <http://www.quantpsy.org/sobel/sobel.htm>

$$z - value = \frac{a*b}{\sqrt{b^2*S_a^2 + a^2*S_b^2}}$$

where :



- a is the unstandardised Coefficient of the independent variable
- S_a is the standard error of the independent variable
- b is the unstandardised Coefficient of the mediation variable
- S_b is the standard error of the mediation variable

Downsides, the distribution of $a*b$ is only normal for large samples, therefore people use a bootstrap method to estimate the confidence interval of ab . If it comprises 0 it is not significant, else it is.

```
# a path
```

```
coef(summary(a.path))
```

```
##           Estimate Std. Error  t value    Pr(>|t|)
## (Intercept) 11.568627  0.5344298 21.646673 4.337682e-54
## conditionPSI 6.819128 0.7634712 8.931742 2.902753e-16
```

a is the unstandardised coefficient for the path from X to M

a = 6.819128

Sa = 0.7634712

```
# c_dash path
```

```
coef(summary(c.dash.path))
```

```
##           Estimate Std. Error  t value    Pr(>|t|)
## (Intercept) -1.4573040 0.17336240 -8.406113 8.413584e-15
## conditionPSI 0.4763879 0.15987428 2.979766 3.247729e-03
## solutions 0.1032689 0.01256431 8.219223 2.700434e-14
```

b refers to the unstandardised coefficient for the path from M to Y

b=0.1032689

Sb=0.01256431

```
# Average Causal Mediation Effect (ACME)
```

```
acme <- a*b
```

```
acme
```

```
## [1] 0.7042038
```

```
# Sobel's Z
```

```
z <- (a*b) / sqrt(b^2 * Sa^2 + b^2 * Sa^2)
```

```
z
```

```
## [1] 6.315696
```

Doing the z-test

We now determine whether this mediation effect is statistically significant (not due to sampling error) with a two-tailed z-test of the hypothesis that the mediated effect equals zero in the population.

H_0 : the mediation effect is zero

H_1 : the mediation effect is not zero

```
# The critical value (1.96 for a two tailed test with alpha = 0.05)
critical.value <- qnorm(0.025, mean = 0, sd = 1, lower.tail = FALSE)
critical.value
```

```
## [1] 1.959964
```

The z-value we obtained (6.33) is much higher than the critical value of 1.96.

We can compute the p-value associated with it. Since it is much lower than alpha (0.05) we can reject H_0 and conclude that the mediation effect is significant.

```
p.value <- pnorm(q = z, mean = 0, sd = 1, lower.tail = FALSE)
p.value
```

```
## [1] 1.344745e-10
```

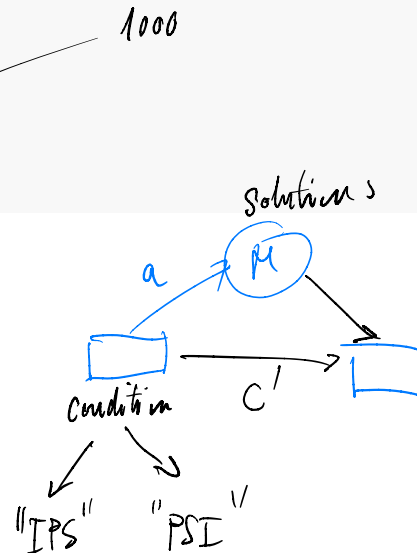
≈ 0

Estimating the proportion of mediation with the mediation package

To alleviate the downsides of the z test (not normally distributed for small samples), we can use a bootstrapping method and simulate 1000 samples to estimate the confidence intervals for the indirect and direct effects.

```
library(mediation) # Gives mediate()

results <- mediate(a.path, c.dash.path,
  treat='condition',
  mediator='solutions',
  boot=TRUE, sims=100,
  control.value = "IPS",
  treat.value = "PSI")
```



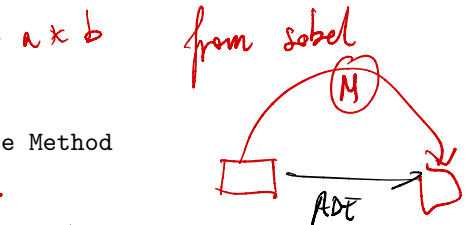
$$c' = \ln(dr \sim iv + m)$$

$$a = \ln(m \sim iv)$$

The direct effect (c path) is listed as ADE (average direct effect), the mediation path (ab path) is listed as ACME (average causal mediation effects, ACME). The total effect is ACE + ACME.

```
summary(results)
```

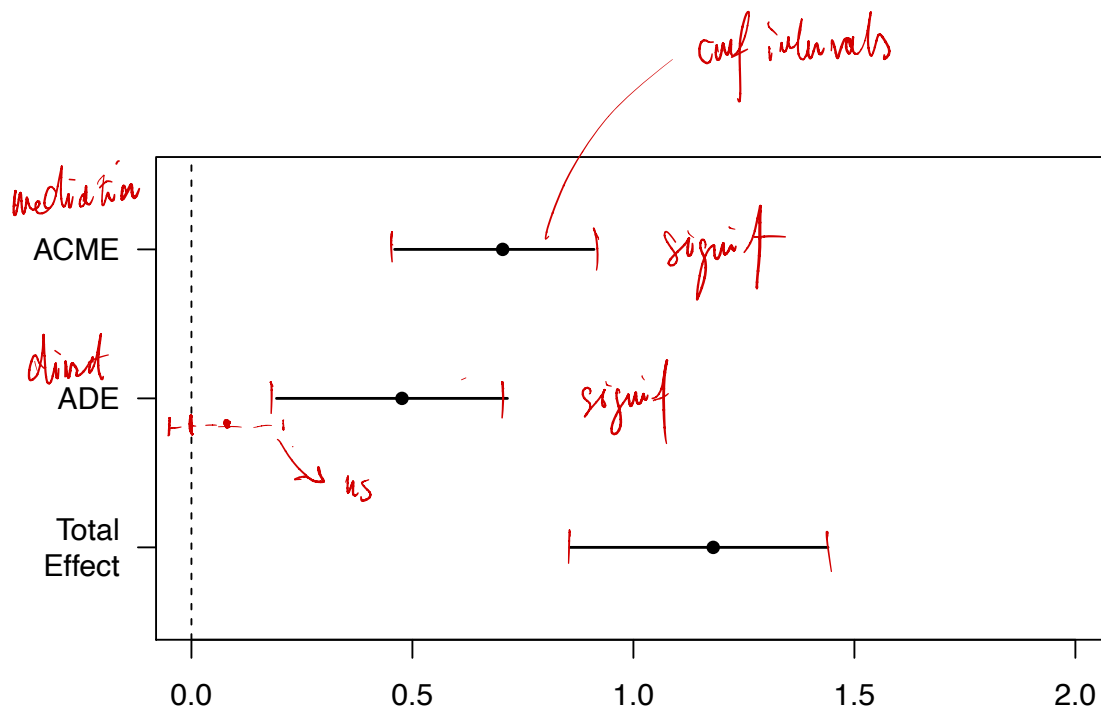
```
##
## Causal Mediation Analysis
##
## Nonparametric Bootstrap Confidence Intervals with the Percentile Method
##
##           Estimate 95% CI Lower 95% CI Upper p-value
## ACME           0.704      0.460      0.91 <2e-16 ***
## ADE            0.476      0.192      0.72  0.02 *
## Total Effect   1.181      0.856      1.44 <2e-16 ***
## Prop. Mediated  0.596      0.423      0.83 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Sample Size Used: 200
##
##
## Simulations: 100
```



stick there

no mediation
ACME ns
ADE *

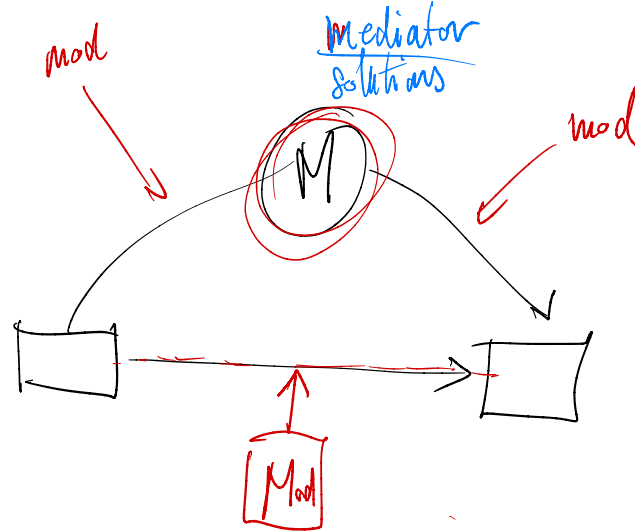
total mediation
ACME *
ADE n.s.
partial med
ACME *
ADE *



Moderated mediation

Remember we found out in the previous ANOVA analyses that the age group was a moderator for the effect of the experimental treatment.

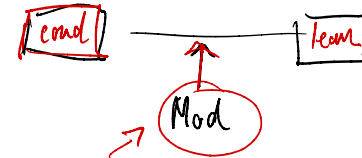
We now look into whether the mediation is also moderated by this variable.



moderator : age group

is the mediation by solutions the same for different age groups

M



Model c

Predicting the dependent variable with the condition

NB: this was our model.2 when doing 2-factor ANOVA earlier. In this context we use the default contrasts (“treatment”) when building the lm model.

```
model.c <- lm(learning ~ condition + age.group + condition:age.group, data=df)
coef(summary(model.c))
```

	Estimate	Std. Error	t value
## (Intercept)	-0.6108569	0.1802479	-3.388983
## conditionPSI	1.1814799	0.2731392	4.325559
## age.groupstudents	0.2800331	0.2549090	1.098561
## age.groupprofessionals	0.8037029	0.2608150	3.081506
## conditionPSI:age.groupstudents	0.4406413	0.3701604	1.190406
## conditionPSI:age.groupprofessionals	-0.6020656	0.3802612	-1.583295
##	Pr(> t)		
## (Intercept)	8.496561e-04		
## conditionPSI	2.432415e-05		
## age.groupstudents	2.733210e-01		
## age.groupprofessionals	2.359029e-03		
## conditionPSI:age.groupstudents	2.353413e-01		
## conditionPSI:age.groupprofessionals	1.149831e-01		

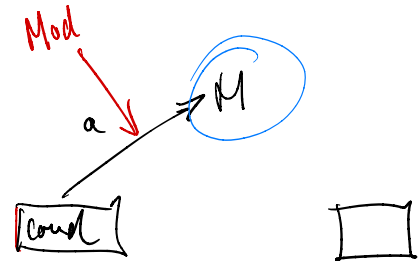
Model a

Predicting the mediator variable with the condition

```
model.a <- lm(solutions ~ condition + age.group + condition:age.group, data=df)
coef(summary(model.a))
```

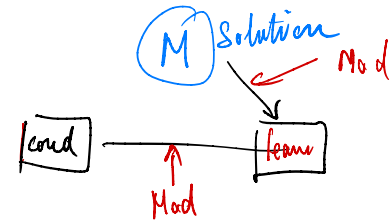
	Estimate	Std. Error	t value	Pr(> t)
## (Intercept)	8.428571	0.6949723	12.127925	1.466060e-25
## conditionPSI	3.164021	1.0531287	3.004401	3.012129e-03
## age.groupstudents	2.971429	0.9828392	3.023311	2.838132e-03
## age.groupprofessionals	6.758929	1.0056107	6.721218	1.957578e-10
## conditionPSI:age.groupstudents	6.593874	1.4272079	4.620121	6.978416e-06
## conditionPSI:age.groupprofessionals	2.406055	1.4661531	1.641066	1.024039e-01

$p < .05$



Model c.dash

Same as model c but we add the mediator (solutions).



```
model.c.dash <- lm(learning ~ condition + age.group + condition:age.group +
  solutions,
  data=df) + solution: age group
coef(summary(model.c.dash))
```

	Estimate	Std. Error	t value
## (Intercept)	-1.67136590	0.20953421	-7.9765777
## conditionPSI	0.78337292	0.24496999	3.1978322
## age.groupstudents	-0.09384122	0.22868401	-0.4103532
## age.groupprofessionals	-0.04672643	0.25388833	-0.1840432
## solutions	0.12582310	0.01632507	7.7073524
## conditionPSI:age.groupstudents	-0.38902032	0.34190884	-1.1377896
## conditionPSI:age.groupprofessionals	-0.90480279	0.33568276	-2.6954104
## Pr(> t)			
## (Intercept)	1.299331e-13		
## conditionPSI	1.618355e-03		
## age.groupstudents	6.820020e-01		
## age.groupprofessionals	8.541728e-01		
## solutions	6.605332e-13		
## conditionPSI:age.groupstudents	2.566188e-01		
## conditionPSI:age.groupprofessionals	7.650840e-03		

With mediate() package

```
results <- mediate(model.a, # predicts mediator with condition  
                  model.c.dash, # predicts learning with condition and solutions  
                  treat='condition',  
                  mediator='solutions',  
                  boot=TRUE, sims=100, # run bootstrapping procedure  
                  control.value = "IPS",  
                  treat.value = "PSI")
```

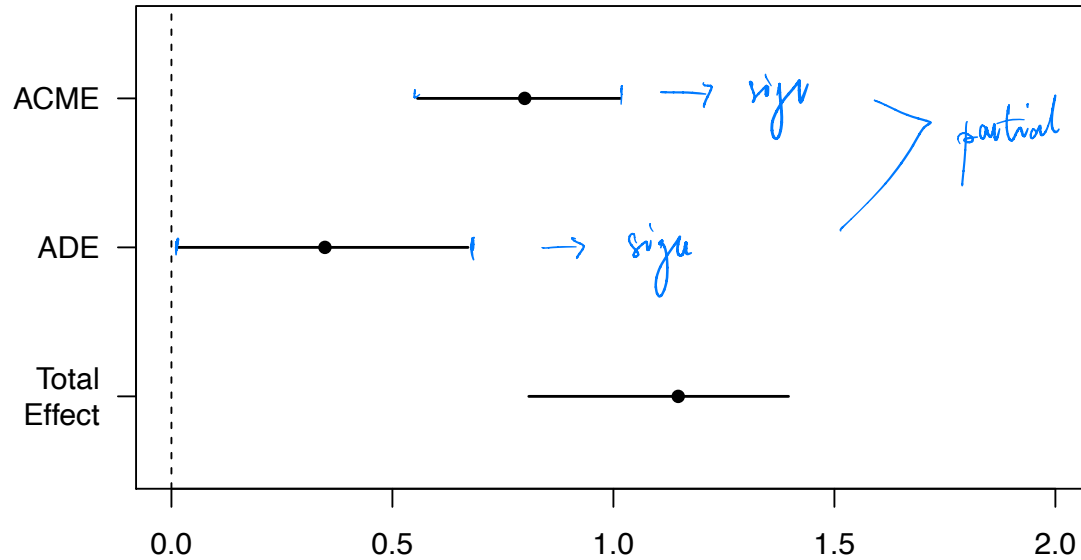
The results indicate a **partial** mediation (ACME **and** ADE are significant).

```
summary(results)
```

```
##
## Causal Mediation Analysis
##
## Nonparametric Bootstrap Confidence Intervals with the Percentile Method
##
##           Estimate 95% CI Lower 95% CI Upper p-value
## ACME           0.7993      0.5568      1.02 <2e-16 ***
## ADE            0.3473      0.0109      0.67  0.04 *
## Total Effect   1.1466      0.8086      1.40 <2e-16 ***
## Prop. Mediated 0.6971      0.4935      0.99 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Sample Size Used: 200
##
##
## Simulations: 100
```

partial

Partial mediation

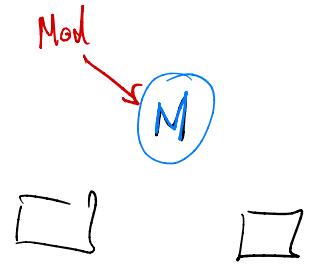
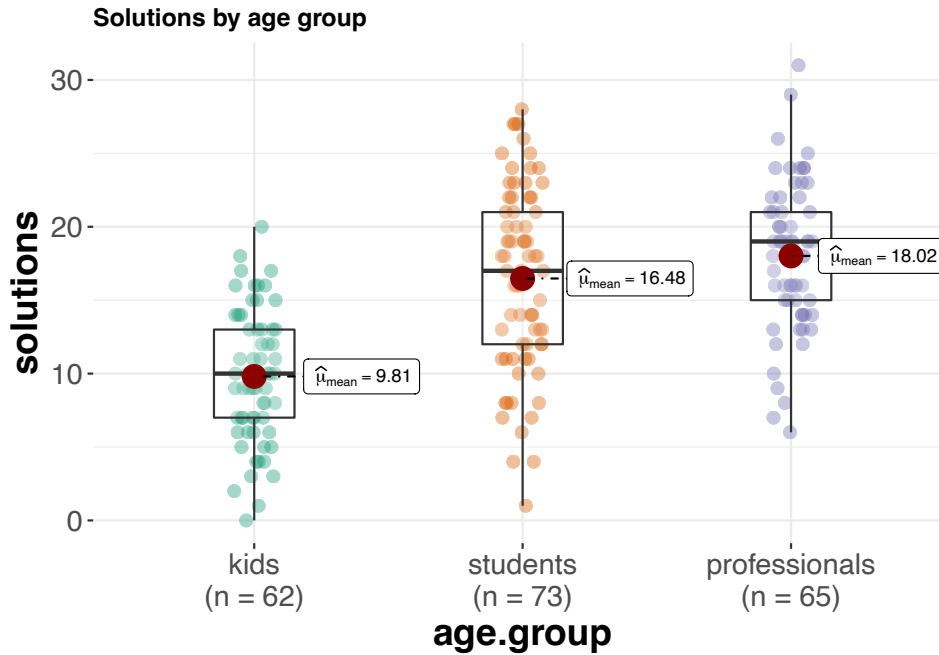


Moderated mediation with age group

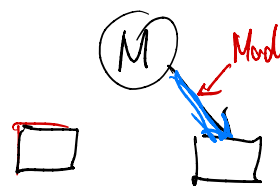
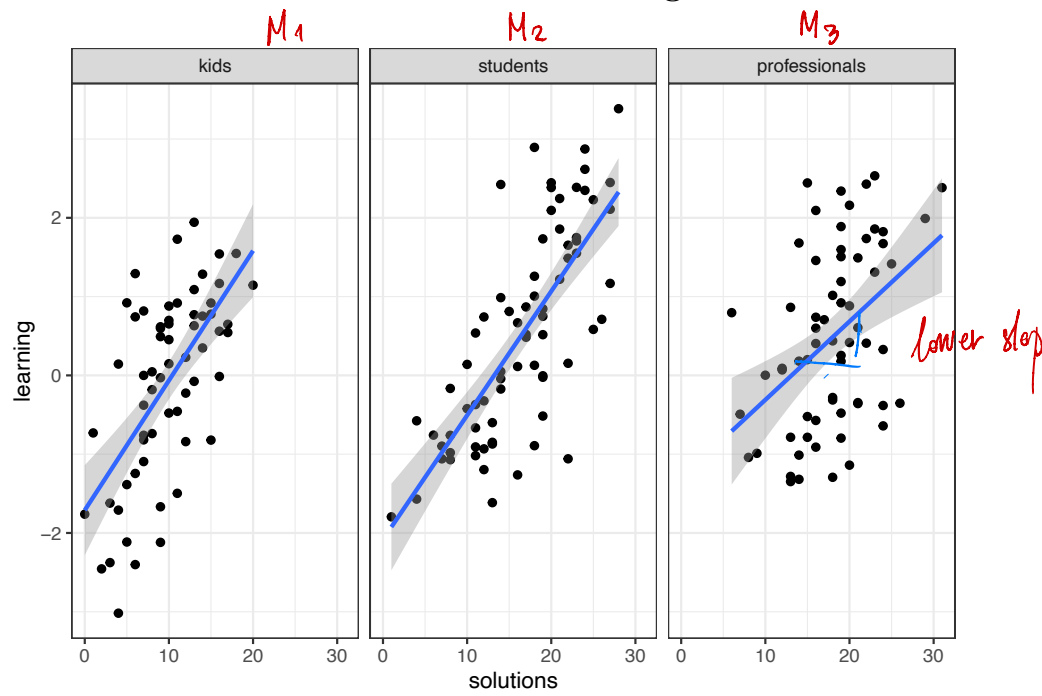
- It appears that kids have created less solutions than students and professionals.
- It appears that the relation between solutions and learning is different for different age groups.

Does the mediation exist for all age groups ?

kids have created less solutions than students and professionals



the relation between solutions and learning is different for different age groups



Adding covariates to the `mediate()` function

Approach: We compute the mediation for each age subgroup. This is done by adding a `covariates` argument to the `mediate` function.

```
results.kids <- mediate(model.a, model.c.dash, treat='condition', mediator='solutions', boot=TRUE, sims=
  covariates = list(age.group="kids"))
summary(results.kids)
```

```
##
## Causal Mediation Analysis
##
## Nonparametric Bootstrap Confidence Intervals with the Percentile Method
##
## (Inference Conditional on the Covariate Values Specified in 'covariates')
##
##      Estimate 95% CI Lower 95% CI Upper p-value
## ACME          0.398      0.191      0.71 <2e-16 ***
## ADE           0.783      0.406      1.17 <2e-16 ***
## Total Effect   1.181      0.658      1.58 <2e-16 ***
## Prop. Mediated 0.337      0.177      0.55 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Sample Size Used: 200
##
##
## Simulations: 100
```

```
results.students <-mediate(model.a, model.c.dash, treat='condition', mediator='solutions', boot=TRUE,si
                           covariates = list(age.group="students"))
summary(results.students)
```

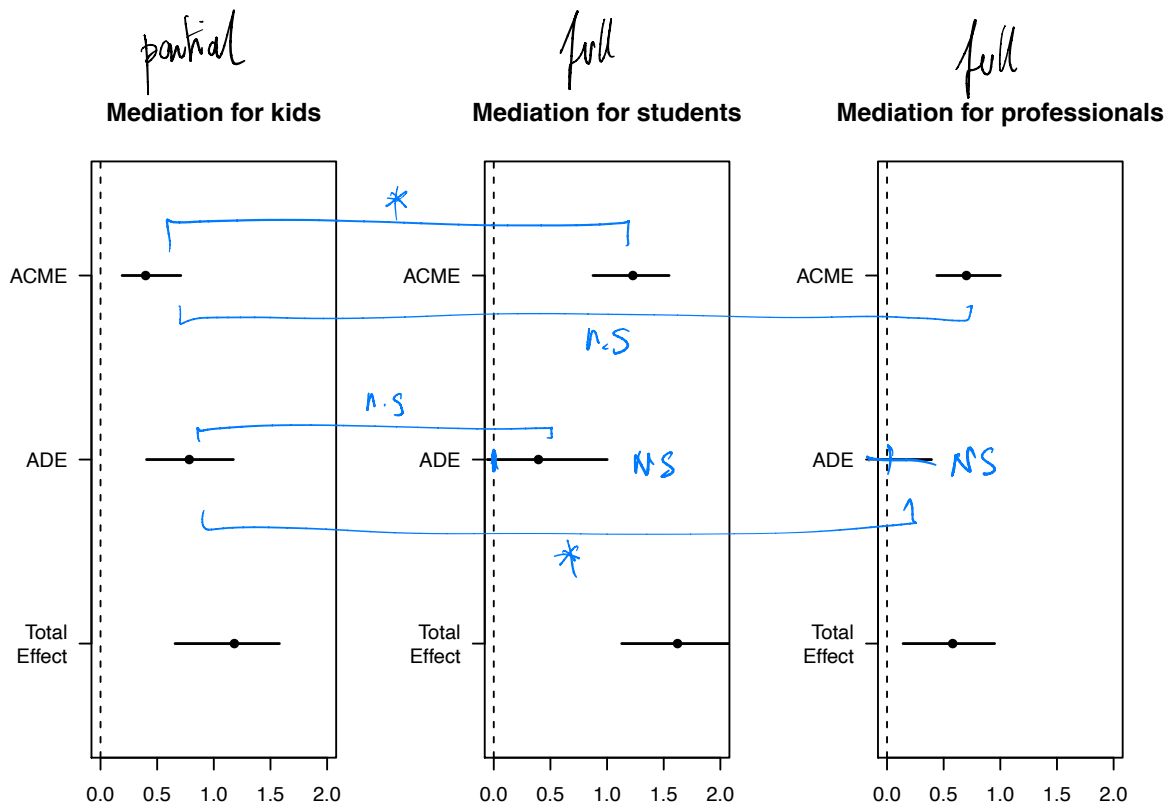
```
##
## Causal Mediation Analysis
##
## Nonparametric Bootstrap Confidence Intervals with the Percentile Method
##
## (Inference Conditional on the Covariate Values Specified in 'covariates')
##
##           Estimate 95% CI Lower 95% CI Upper p-value
## ACME           1.2278      0.8760      1.55 <2e-16 ***
## ADE             0.3944     -0.0568      1.00   0.12
## Total Effect    1.6221      1.1293      2.09 <2e-16 ***
## Prop. Mediated   0.7569      0.5137      1.05 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Sample Size Used: 200
##
##
## Simulations: 100
```

no direct effect
→ Full mediation

```
results.professionals <- mediate(model.a, model.c.dash, treat='condition', mediator='solutions', boot=T
                                covariates = list(age.group="professionals"))
summary(results.professionals)
```

```
##
## Causal Mediation Analysis
##
## Nonparametric Bootstrap Confidence Intervals with the Percentile Method
##
## (Inference Conditional on the Covariate Values Specified in 'covariates')
##
##           Estimate 95% CI Lower 95% CI Upper p-value
## ACME           0.701      0.440      1.00 <2e-16 ***
## ADE            -0.121     -0.601      0.39   0.74
## Total Effect    0.579      0.142      0.95 <2e-16 ***
## Prop. Mediated  1.210      0.562      4.56 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Sample Size Used: 200
##
##
## Simulations: 100
```

no more direct

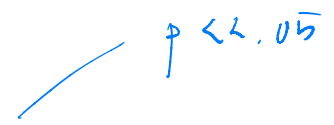


Pushing the envelope ...

Testing whether the direct and indirect effects are different across groups ?

```
test.modmed(results,  
  covariates.1 = list(age.group = "kids"),  
  covariates.2 = list(age.group = "students"), sims = 100)
```

```
##  
## Test of  $ACME(covariates.1) - ACME(covariates.2) = 0$   
##  
## data: estimates from results  
##  $ACME(covariates.1) - ACME(covariates.2) = -0.82966$ , p-value <  $2.2e-16$   
## alternative hypothesis: true  $ACME(covariates.1) - ACME(covariates.2)$  is not equal to 0  
## 95 percent confidence interval:  
## -1.330910 -0.344623  
##  
##  
## Test of  $ADE(covariates.1) - ADE(covariates.2) = 0$   
##  
## data: estimates from results  
##  $ADE(covariates.1) - ADE(covariates.2) = 0.38902$ , p-value = 0.34  
## alternative hypothesis: true  $ADE(covariates.1) - ADE(covariates.2)$  is not equal to 0  
## 95 percent confidence interval:  
## -0.160029 1.071236
```



```

test.modmed(results,
             covariates.1 = list(age.group = "kids"),
             covariates.2 = list(age.group = "professionals"), sims = 100)

##
## Test of ACME(covariates.1) - ACME(covariates.2) = 0
##
## data: estimates from results
## ACME(covariates.1) - ACME(covariates.2) = -0.30274, p-value = 0.1
## alternative hypothesis: true ACME(covariates.1) - ACME(covariates.2) is not equal to 0
## 95 percent confidence interval:
## -0.71011597 0.02555824
##
##
## Test of ADE(covariates.1) - ADE(covariates.2) = 0
##
## data: estimates from results
## ADE(covariates.1) - ADE(covariates.2) = 0.9048, p-value < 2.2e-16
## alternative hypothesis: true ADE(covariates.1) - ADE(covariates.2) is not equal to 0
## 95 percent confidence interval:
## 0.3385389 1.7701618

```

```

test.modmed(results,
             covariates.1 = list(age.group = "students"),
             covariates.2 = list(age.group = "professionals"), sims = 100)

##
## Test of ACME(covariates.1) - ACME(covariates.2) = 0
##
## data: estimates from results
## ACME(covariates.1) - ACME(covariates.2) = 0.52692, p-value = 0.04
## alternative hypothesis: true ACME(covariates.1) - ACME(covariates.2) is not equal to 0
## 95 percent confidence interval:
## 0.05270736 1.07853506
##
##
## Test of ADE(covariates.1) - ADE(covariates.2) = 0
##
## data: estimates from results
## ADE(covariates.1) - ADE(covariates.2) = 0.51578, p-value = 0.1
## alternative hypothesis: true ADE(covariates.1) - ADE(covariates.2) is not equal to 0
## 95 percent confidence interval:
## -0.1256427 1.2747345

```

Role of a pre-treatment moderator ?

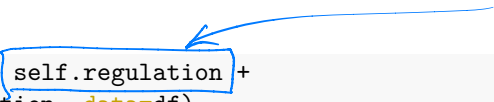
We now look at the potential influence of a potential pre-treatment moderator, the `self.regulation` skills of the participants.

We wonder whether the potential positive effect of testing many solutions is conditioned on the level of self-regulation skills.

The rationale for this could be: when you run experiments (generate many `solutions`), you learn more if you can accurately revise your hypotheses (`self-regulation`) about the phenomenon. This would mean that the moderation effect that we discovered for the number of solutions would be present mainly for subjects with a high level of self-regulation skills.

The analysis we conduct is similar to the one we did for `age.group` except that this time, the moderator is a continuous variable.

```
model.a2 <- lm(solutions ~ condition + self.regulation +  
               condition:self.regulation, data=df)  
  
model.c2.dash <- lm(learning ~ condition + self.regulation +  
                   condition:self.regulation +  
                   solutions, data=df)
```



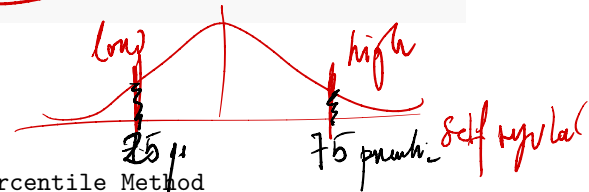
condition's

```
results.self.regulation <- mediate(model.a2, model.c2.dash, treat='condition', mediator='solutions',
                                boot=TRUE,sims=100, control.value = "IPS", treat.value = "PSI")
summary(results.self.regulation)
```

```
##
## Causal Mediation Analysis
##
## Nonparametric Bootstrap Confidence Intervals with the Percentile Method
##
##           Estimate 95% CI Lower 95% CI Upper p-value
## ACME           0.478      0.286      0.70 <2e-16 ***
## ADE            0.473      0.226      0.79 <2e-16 ***
## Total Effect    0.950      0.699      1.22 <2e-16 ***
## Prop. Mediated   0.503      0.336      0.71 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Sample Size Used: 200
##
##
## Simulations: 100
```

```
results.hi.self.regulation <- mediate(model.a2, model.c2.dash, treat='condition', mediator='solutions',
                                     boot=TRUE, sims=500, control.value = "IPS", treat.value = "PSI",
                                     covariates = list(self.regulation=quantile(df$self.regulation, .75)))
summary(results.hi.self.regulation)
```

```
##
## Causal Mediation Analysis
##
## Nonparametric Bootstrap Confidence Intervals with the Percentile Method
##
## (Inference Conditional on the Covariate Values Specified in 'covariates')
##
##      Estimate 95% CI Lower 95% CI Upper p-value
## ACME          0.4700      0.2772      0.71 <2e-16 ***
## ADE           0.3437     -0.0347      0.70  0.088 .
## Total Effect   0.8137      0.4119      1.21 <2e-16 ***
## Prop. Mediated 0.5776      0.3458      1.07 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Sample Size Used: 200
##
##
## Simulations: 500
```



```
results.lo.self.regulation <- mediate(model.a2, model.c2.dash, treat='condition', mediator='solutions',
                                     boot=TRUE,sims=500, control.value = "IPS", treat.value = "PSI",
                                     covariates = list(self.regulation=quantile(df$self.regulation, .25)))
summary(results.lo.self.regulation)
```

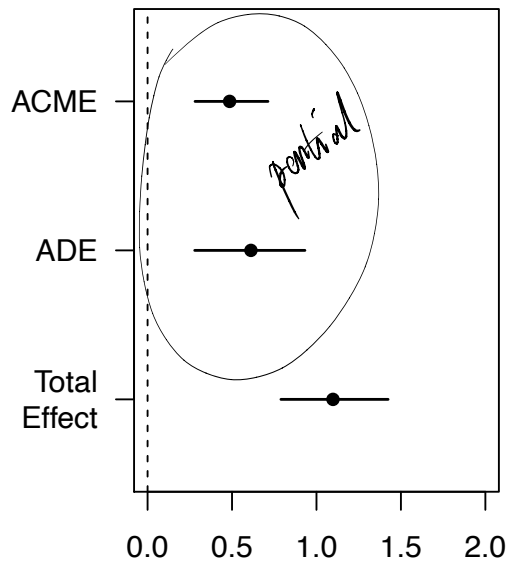
```
##
## Causal Mediation Analysis
##
## Nonparametric Bootstrap Confidence Intervals with the Percentile Method
##
## (Inference Conditional on the Covariate Values Specified in 'covariates')
```

	Estimate	95% CI Lower	95% CI Upper	p-value
## ACME	0.486	0.281	0.71	<2e-16 ***
## ADE	0.612	0.279	0.93	<2e-16 ***
## Total Effect	1.098	0.790	1.42	<2e-16 ***
## Prop. Mediated	0.443	0.271	0.68	<2e-16 ***

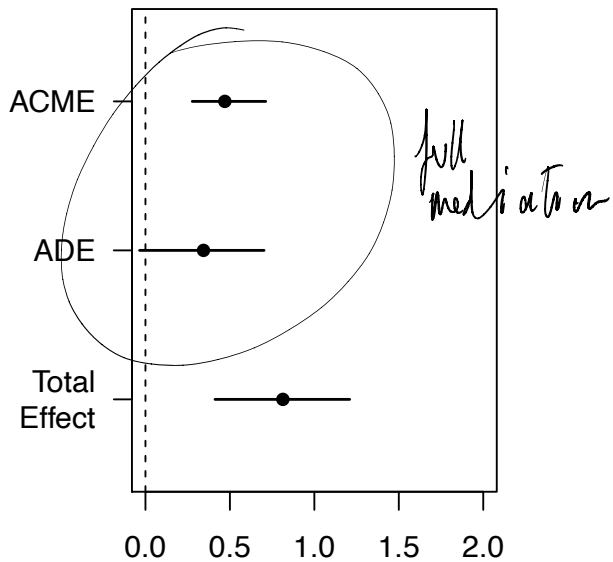
```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Sample Size Used: 200
##
##
## Simulations: 500
```

partial

Low self-regulation



High self-regulation



no moderation : same ADE same ACME for all groups

We finally test whether the mediation and direct effects are different for students with hi and lo self regulation skills

```
test.modmed(results.self.regulation,
             covariates.1 = list(self.regulation=quantile(df$self.regulation, .25)),
             covariates.2 = list(self.regulation=quantile(df$self.regulation, .75)), sims = 500)

##
## Test of ACME(covariates.1) - ACME(covariates.2) = 0
##
## data: estimates from results.self.regulation
## ACME(covariates.1) - ACME(covariates.2) = 0.015756, p-value = 0.88
## alternative hypothesis: true ACME(covariates.1) - ACME(covariates.2) is not equal to 0
## 95 percent confidence interval:
## -0.3241569 0.3702842
##
##
## Test of ADE(covariates.1) - ADE(covariates.2) = 0
##
## data: estimates from results.self.regulation
## ADE(covariates.1) - ADE(covariates.2) = 0.26823, p-value = 0.304
## alternative hypothesis: true ADE(covariates.1) - ADE(covariates.2) is not equal to 0
## 95 percent confidence interval:
## -0.2742372 0.7459550
```