MEDIATION

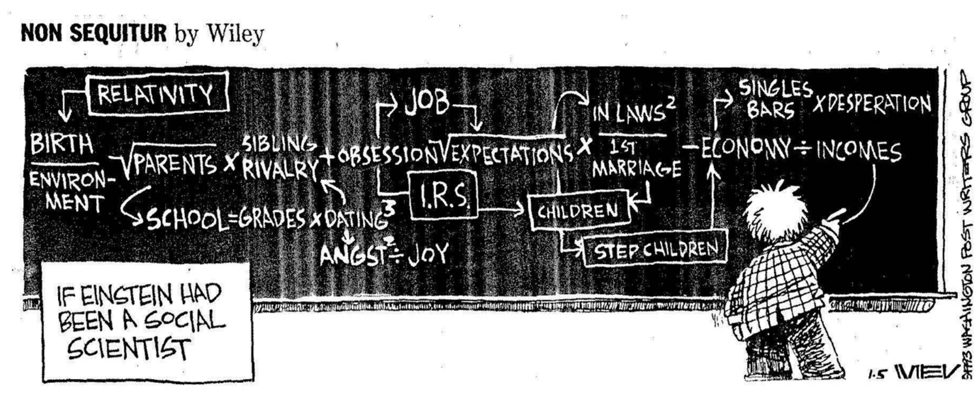
Research Methods in Psychology I & II ▪ Department of Psychology ▪ Colorado State University

BY THE END OF THIS UNIT YOU WILL:

1. Be able to define mediation in the context of regression modeling.
2. Understand the utility of mediation modeling for substantive research.
3. Know how to fit a mediation model in R.
4. Be able to estimate an indirect effect and obtain appropriate confidence intervals.
5. Be familiar with some advanced mediation techniques, such as models with multiple mediators, with dummy coded x variables, and serial mediation.
6. Know how to combine moderation and mediation to consider moderated mediation models.
7. Know how to describe the results of a mediation model.

What is a Mediation Model?

Mediation models help us to examine causal chains or processes. For example, we might be interested in the mechanisms by which a worksite training program improves employee productivity or a new therapy reduces depressive symptoms. These sorts of research questions posit a causal chain in which an antecedent variable (e.g., a new therapy) affects a mediating variable (e.g., coping skills), which in turn affects an outcome variable (e.g., depression).



Overview

What is mediation in the context of regression?

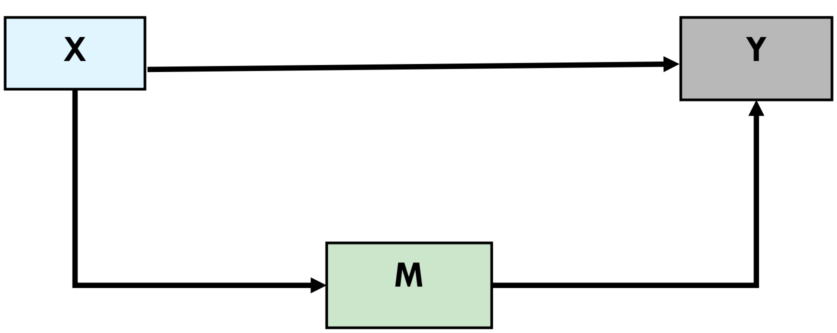
Mediation models help us to examine causal chains or processes. In this context, the effect of some antecedent variable (x) on some outcome (y) is in part (or whole) due to a change in some intermediate variable (i.e., mediator—m). In other words, x causes m, and m in turn causes y.

What does mediation analysis allow us to do?

- Determine the mechanisms by which x causes y.
- In terms of intervention studies—mediation analysis helps us to understand why or how a program affects the desired outcomes.
- In terms of experimental studies—mediation analysis helps us to understand how some manipulation leads to a certain outcome.
- In terms of observational studies—mediation analysis helps us to understand how one observed outcome causes some subsequent observed outcome. Having the right order (i.e., x causes m, and m causes y) and potential confounders measured is critical. If this can't be established then you can only determine if a model is consistent with mediation, not that mediation truly exists.

What are some examples of research questions we can answer?

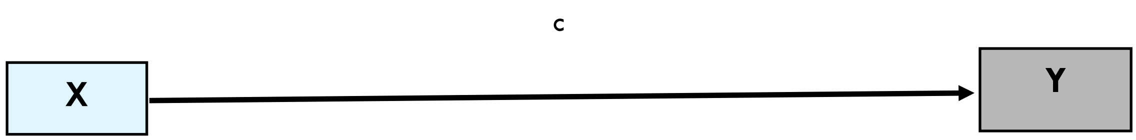
- Counseling Psychology: Does a new therapy positively impact depression by reducing rumination?
- I-O Psychology: Does leadership training for managers result in more productive employees because managers become better communicators?
- Social and Health Psychology: Does a sexual health intervention reduce risky sex by increasing perceived susceptibility to sexually transmitted diseases?
- Cognitive Neuroscience: Do structural variations in the prefrontal cortex mediate the relationship between adolescent stress and psychological health?
- Cognitive Psychology: Does processing speed mediate the effect of age on decision making capacity?



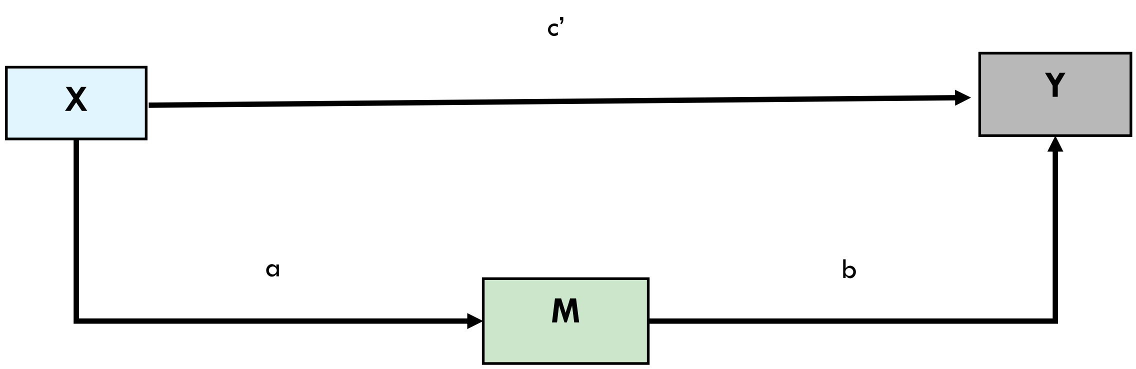
Mediation is a causal model, it is posited that x causes m, and in turn, m causes y. The initial variable is presumed to be causally prior to the mediator and the mediator is presumed to be causally prior to the outcome.

Terminology

Consider a variable (x) that causes another variable (y). The variable x is called the initial or antecedent variable and y is called the outcome. In diagrammatic form, the unmediated model is (where c is the total effect):



The effect of x on y may be explained by a mediating variable m (also called an intermediate variable). The path labeled *a* represents the effect of the antecedent variable on the mediator and the path labeled *b* represents the effect of the mediator (m) on the outcome, after adjusting for the antecedent. **<u>Both of these effects (i.e., paths *a* and *b*) must be statistically significant for mediation to exist.</u>** The variable x may still affect y directly (i.e., partial mediation) or it may not (i.e., full mediation). The mediated model is (where c' [c-prime] is the direct effect:



While c' captures the direct effect (i.e., the effect of x on y that **does not** go through m), the indirect effect of x on y that does go through m is calculated as the product of the "*a*-path" and the "*b*-path"—that is *a*•*b*.

In this way, the total effect (c) is partitioned into the direct effect (c') and the indirect effect (*ab*).

Therefore, when m and y are continuous variables (which is the type of models considered in this unit), we know that:

$c = c' + ab$

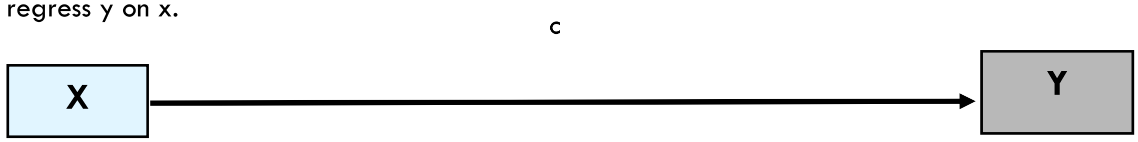
and

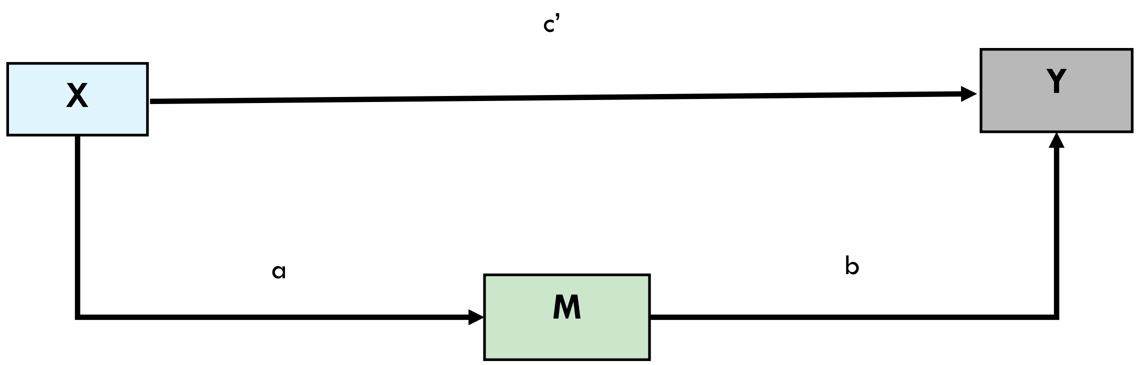
$c - c' = ab$

Estimation

We can estimate each of the effects in the mediation figure with a series of 3 regression models.

MODEL 1: To obtain path c,
we regress y on x.



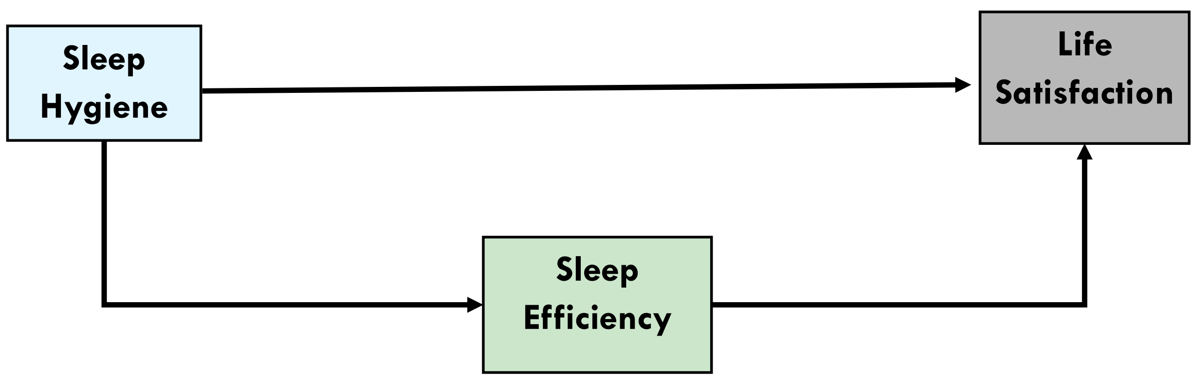


MODEL 2: To obtain path *a*,
we regress m on x.

MODEL 3: To obtain paths c’ and *b*,
we regress y on m and x.

Does Sleep Efficiency Mediate the Effect of Sleep Hygiene on Life Satisfaction?

Recall the sleep data from last semester. Imagine that the researcher sought to determine if the effect of sleep hygiene on life satisfaction was mediated (i.e., explained) by sleep efficiency. In other words, the researcher desired to determine if participants with better sleep hygiene had better sleep efficiency, which in turn led to better life satisfaction. Put another way, the researcher asked: "Does sleep efficiency explain, at least in part, the influence that sleep hygiene has on life satisfaction?" **To keep things simple at first, we will consider just the males in the sample.**



Prepare the data

Let's create a new notebook called: Sleep_Notebook_Mediation. Create the following code chunks.

Load libraries

```
library(tidyverse)
library(olsrr)
library(modelr)
library(huxtable)
library(GGally)
library(car)

library(RMediation)
```

Import data

```
slp <- read_csv("slpdata.csv")
```

Do some formatting

```
slp <- read_csv("slpdata.csv")

slp <- mutate(slp,
        female = ifelse(sex == 1, 0, 1),
        female.f = factor(female, levels = c(0,1), labels = c("male", "female")),
        cond2 = ifelse(cond == 2, 1, 0),
        cond3 = ifelse(cond == 3, 1, 0),
        cond.f = factor(cond, levels = c(1,2,3), labels = c("self help", "group-based", "group + partner")))

# subset to keep males only
slp_males <- filter(slp, female == 0)

# subset to keep females only
slp_females <- filter(slp, female == 1)
```

Estimation

We can estimate each of the effects in the mediation figure with a series of 3 regression models.

Estimate the base models

model.xy

```
model.xy <- lm(lifesat ~ hygiene, data=slp_males)
ols_regress(model.xy)
```

model.xm

```
model.xm <- lm(sleep ~ hygiene , data=slp_males)
ols_regress(model.xm)
```

model.xmy

```
model.xmy <- lm(lifesat ~ sleep + hygiene, data=slp_males)
ols_regress(model.xmy)
```

Output from 3 Regression Models

model.xy

Parameter Estimates

model	Beta	Std. Error	Std. Beta	t	Sig	lower	upper
(Intercept)	2.399	0.161		14.865	0.000	2.082	2.716
hygiene	0.295	0.028	0.486	10.406	0.000	0.239	0.350

This model assesses the c-path. The effect is relatively large and statistically significant. A one unit increase in sleep hygiene is associated with a .295 unit increase in life satisfaction. That is, people with better sleep hygiene tend to have more life satisfaction. This is the total effect of hygiene on life satisfaction.

model.xm

Parameter Estimates

model	Beta	Std. Error	Std. Beta	t	Sig	lower	upper
(Intercept)	35.235	1.802		19.556	0.000	31.691	38.779
hygiene	5.327	0.316	0.669	16.853	0.000	4.705	5.948

This model assesses the *a*-path in the mediation model. It is large and statistically significant, indicating that a one unit increase in sleep hygiene is associated with a 5.327 unit increase in sleep efficiency. People with better sleep hygiene tend to have more efficient sleep.

model.xmy

Parameter Estimates

model	Beta	Std. Error	Std. Beta	t	Sig	lower	upper
(Intercept)	0.959	0.208		4.613	0.000	0.550	1.368
sleep	0.041	0.004	0.537	9.594	0.000	0.032	0.049
hygiene	0.077	0.034	0.127	2.266	0.024	0.010	0.144

This model assesses the *b*- and c'-paths in the mediation model. The *b*-path is large and statistically significant, indicating that, holding constant sleep hygiene, a one unit increase in sleep efficiency is associated with a .04 unit increase in life satisfaction. The c'-path is small but also statistically significant, indicating that holding constant sleep efficiency, a one unit increase in sleep hygiene is associated with a .08 unit increase in life satisfaction. Since c' remains significant, the mediation is only partial at best. That is, a direct effect of sleep hygiene on life satisfaction remains after accounting for sleep efficiency.

Anatomy of a Regression Model Object

In this unit, we're going to pull out various estimates from our regression model objects to use them for the creation of subsequent estimates and significance tests (e.g., of the indirect effect — a-path • b-path). Let's take a moment to see what R stores for us after estimating a linear model and how we access the quantities of interest. Let's consider model.xmy.

coefficients(model.xmy)

This refers to the estimates of the intercept and slopes.

```
(Intercept)       sleep     hygiene
 0.95883514  0.04087233  0.07688158
```

vcov(model.xmy)

This refers to the variance-covariance matrix of the regression model results. On the diagonal of the matrix is the variability of the intercept and slopes. Taking the square root of these values on the diagonal yields the standard error of each estimate.

```
              (Intercept)         sleep       hygiene
(Intercept)  4.321244e-02 -6.394743e-04 -9.405536e-05
sleep       -6.394743e-04  1.814882e-05 -9.667310e-05
hygiene     -9.405536e-05 -9.667310e-05  1.151356e-03
```

path.b <- coefficients(model.xmy)["sleep"]

se.path.b<-sqrt(vcov(model.xmy)["sleep","sleep"])

By naming and storing these values as objects, we can use them later to construct our estimate and significance test of the indirect effect.

```
     sleep
0.04087233

0.004260143
```

Parameter Estimates

model	Beta	Std. Error	Std. Beta	t	Sig	lower	upper
(Intercept)	0.959	0.208		4.613	0.000	0.550	1.368
sleep	0.041	0.004	0.537	9.594	0.000	0.032	0.049
hygiene	0.077	0.034	0.127	2.266	0.024	0.010	0.144

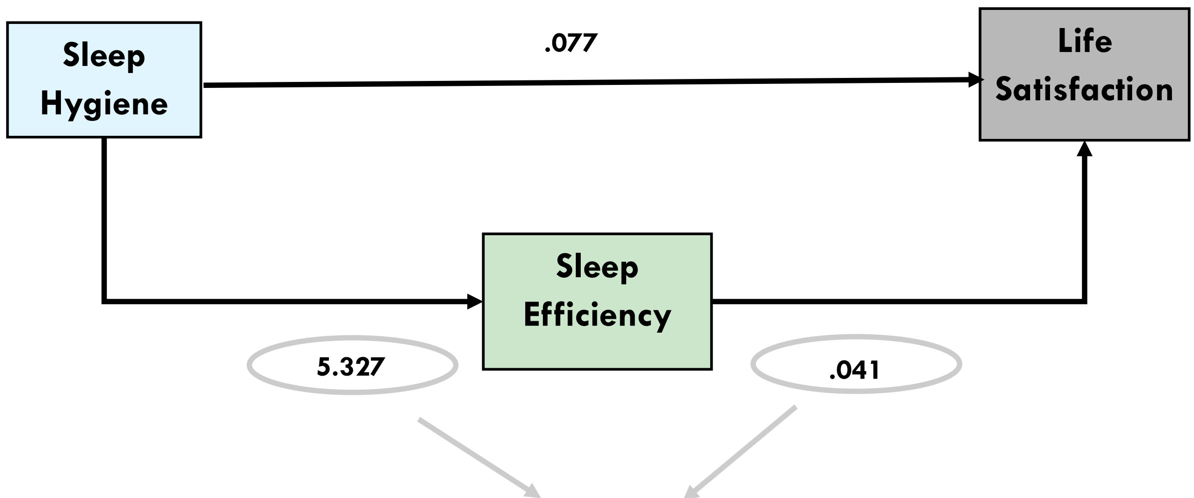
Look at components of the lm object

```
coefficients(model.xm)
coefficients(model.xmy)
vcov(model.xm)
vcov(model.xmy)
```

Pull out and name certain components

```
path.a <- coefficients(model.xm)["hygiene"]
path.b <- coefficients(model.xmy)["sleep"]

se.path.a<-sqrt(vcov(model.xm)["hygiene","hygiene"])
se.path.b<-sqrt(vcov(model.xmy)["sleep","sleep"])
```

Calculate the Indirect Effect



$$a \cdot b = 5.327 \cdot .041 = .218$$

$$c - c' = .295 - .077 = .218$$

Calculate the indirect effect—product term

```
ab <- path.a*path.b
ab
```

We know from our regression model results that both the a-path and the b-path are statistically significant. Therefore, our model meets the minimum criteria for mediation.

Significance Test for Indirect Effect (Confidence Interval)

Now we have an estimate of the indirect effect (.218). We also need to determine if the indirect effect is significantly different from zero. In the 1980's, when tests of mediation were initially proposed, something called a Sobel Test was recommended for testing the significance of an indirect effect. The Sobel Test provided a formula for calculating a standard error, and it was recommended that the indirect effect estimate should be divided by the standard error to obtain something akin to a t^* to determine if the indirect effect was significantly different than 0. When we divide an estimate (e.g., *ab*), by its standard error (se_{ab}), and compare this ratio to a standard normal distribution, we assume that the sampling distribution of the estimate (e.g., *ab*) is normal. However, simulation studies have demonstrated that while the estimates of *a* and *b* are typically normal, the product of the two (*ab*) is often not—rather it is often positively skewed and leptokurtic (acute peak). This compromises our ability to have confidence in the Sobel Test, particularly when the sample size is small. Instead of assuming the sampling distribution of *ab* is normal, there are a few alternative methods to calculate the confidence interval for the indirect effect in which the normality assumption is not made. One is to use the Distribution of the PRODuct Confidence Limits for INdirect effects (PRODCLIN) method proposed by Mackinnon et al. (2007), and implemented in the RMediation package. We will explore that approach first. In addition, one of the best, and most widely used, alternative methods involves the calculation of a bootstrapped confidence interval for the indirect effect (*ab*). We will explore this approach second.

MacKinnon, D.P., Fritz, M.S., Williams, J., & Lockwood, C.M. (2007). Distribution of the product confidence limits for an indirect effect: Program PRODCLIN. *Behavioral Research Methods, 39(3), 384-389.*

PRODCLIN Method for Statistical Inference

Calculate the confidence interval via PRODCLIN

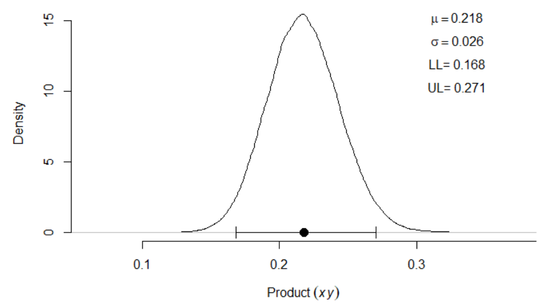
List names of a-path, b-path, standard error of a-path, and standard error of b-path in that order

Desired alpha for construction of CI

These last two arguments ask for a plot of the indirect effect with a CI

```
medci(path.a, path.b, se.path.a, se.path.b, alpha = 0.05, plot = TRUE,  plotCI = TRUE)
```

```
$`97.5% CI`
[1] 0.1681202 0.2705924

$Estimate
  hygiene
0.2177142

$SE
   hygiene
0.02614678
```



The indirect effect is .218, and the associated standard error is .026. The 95% CI for the indirect effect is .168 to .271. Notice that this CI doesn't contain 0, suggesting that 0 is not a plausible value for the indirect effect. Our model.xmy shows that the direct effect (c') is .077 and we have now calculated the indirect effect to be .218. This indicates that part of the total effect of hygiene on life satisfaction DOES NOT go through sleep efficiency (the direct effect, c'-path), and part of the total effect of hygiene on life satisfaction DOES go through sleep efficiency (the indirect effect, ab). We can calculate the proportion of the total effect that is indirect via sleep efficiency as follows: $(a \cdot b)/c = .218/.295 = .74$, indicating 74% of the total effect of sleep hygiene on life satisfaction is indirect through sleep efficiency.

Bootstrapping—An Alternative for Calculating the CI of the Indirect Effect

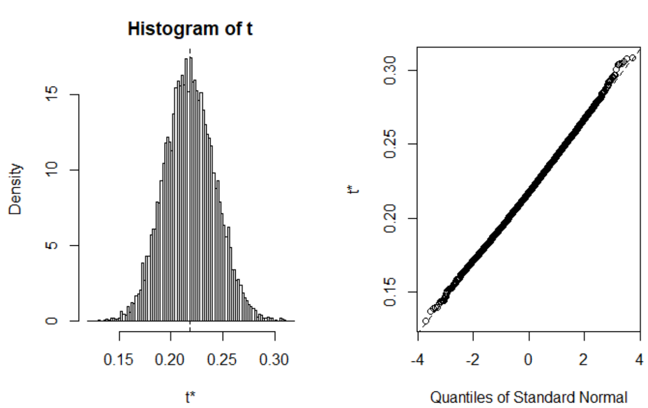
We can also bootstrap the confidence interval using the same techniques we studied in Unit 9. Recall that bootstrapping is recommended when a normal sampling distribution cannot be assumed, which is the worry with an indirect effect.

Calculate the confidence interval via bootstrap

```
boot.med <- function(data, indices){
    data <-data[indices,]

    model.xm <- lm(sleep ~ hygiene , data=data)
    model.xmy <- lm(lifesat ~ sleep + hygiene, data=data)

    path.a <- coefficients(model.xm)["hygiene"]        #the a-path coefficient
    path.b <- coefficients(model.xmy)["sleep"]         #the b-path coefficient
    ab <- path.a*path.b                                #returns the product of a*b
    return(ab)
}

medboot<-boot(data=slp_males, statistic = boot.med, R = 10000)
medconfint <- boot.ci(medboot, index=1, conf = (.95), type = "bca")
print (medconfint)

plot(medboot)
```

The code for this is a little tricky— I underlined the parts that you would change for an alternative example.

```
BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS
Based on 10000 bootstrap replicates

CALL :
boot.ci(boot.out = medboot, conf = (0.95), type = "bca", index = 1)

Intervals :
Level       BCa
95%   ( 0.1742,  0.2682 )
Calculations and Intervals on Original Scale
```



Calling the boot object (named medboot in our syntax above) with the plot function will produce these graphs, which represent an estimate of the indirect effect across all of the bootstrap resamples. You want the density curve to look "filled in" - that is, not sparse. If it looks sparse then you should increase the number of bootstrap resamples (R = 20000). You also want the estimate to be stable, so, for example, the CI is similar when R = 10,000 and R = 11,000.

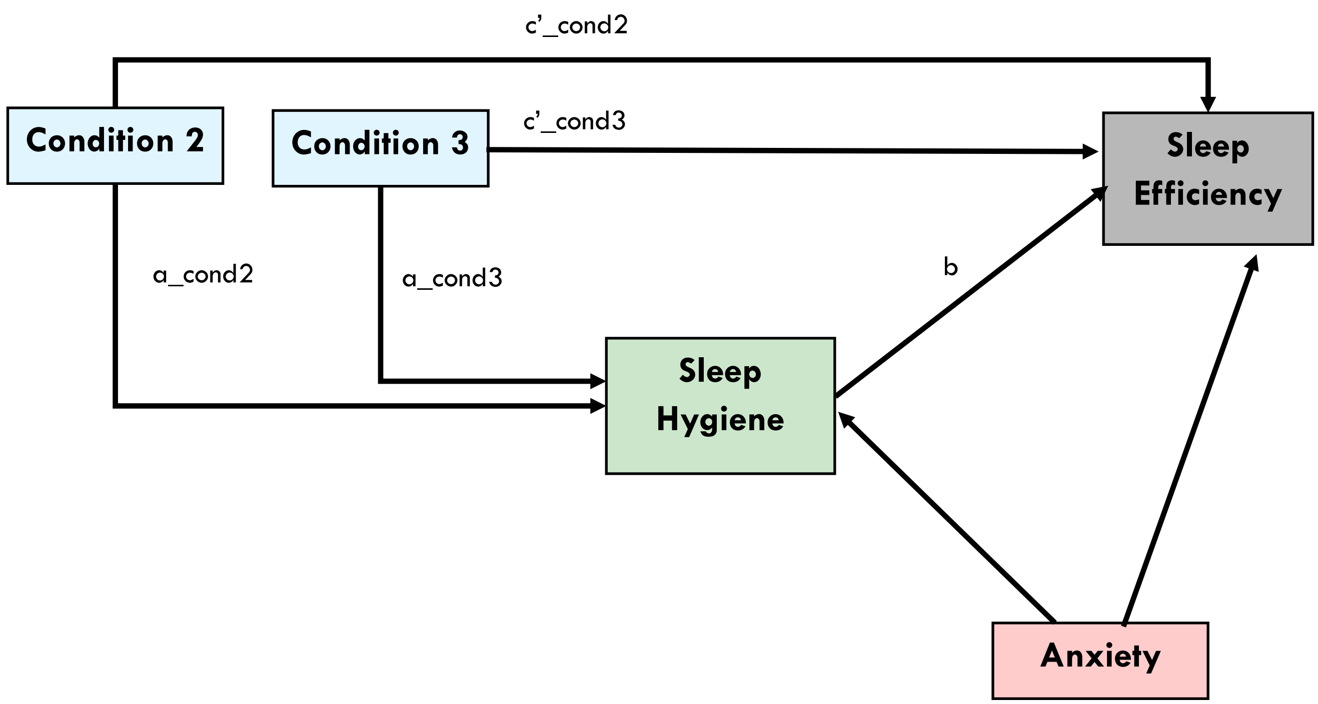
Write up of Results

Sleep efficiency was examined as a potential mediator (m) of the relationship between sleep hygiene (the independent variable—x) and life satisfaction (the outcome—y) among males taking part in a sleep intervention. We hypothesized that better sleep hygiene would lead to better sleep efficiency (i.e., the a-path of the mediation model), which in turn would lead to greater life satisfaction (the b-path of the mediation model). Three regression models were estimated. In the first, life satisfaction (y) was regressed on sleep hygiene (x). In the second, sleep efficiency (m) was regressed on sleep hygiene (x). In the third, life satisfaction (y) was regressed on both sleep hygiene (x) and sleep efficiency (m). Consistent with our hypotheses, sleep hygiene was positively and significantly associated with sleep efficiency (*b*=5.33, 95% CI 4.71, 5.95) and sleep efficiency was positively and significantly associated with life satisfaction (*b*=.04, 95% CI .03, .05), holding constant sleep hygiene. Sleep hygiene was positively and significantly associated with life satisfaction both before (*b*=.30 95% CI .24, 35) and after (*b*=.08, 95% CI .01, .14) the inclusion of the mediator, although its effect was substantially reduced. The indirect effect was estimated to be .22. To determine if the indirect effect was significantly different from zero, we drew 10000 bootstrap samples. The 95% bias-corrected confidence interval for the indirect effect indicated that the effect was significantly different from zero (95% CI .17, .27). Thus, sleep efficiency was a significant partial mediator of the relationship between sleep hygiene and life satisfaction. The ratio of the observed indirect effect (.22), to the observed total effect (.30) is .74, indicating that about 74% of the observed effect of sleep hygiene on life satisfaction was mediated (i.e., explained) by sleep efficiency.

Advanced Example 1: Mediation With Dummy Coded X and Covariates

Let's consider a slightly more complex example. Let's determine if the treatment effects of the group-based (Condition 2) and the group + partner (Condition 3) programs on sleep efficiency can be explained by sleep hygiene. Let's also examine how covariates can be introduced. We will consider anxiety, but any variable that is important to control or adjust for when examining the effects of interest can be introduced. For this example we will consider females only.

The difference between this model and the prior model is that we now have two c-paths, two c'-paths, and two a-paths—that is, one set for the Condition 2 dummy code and one set for the Condition 3 dummy code.



Execute the Models

Estimate the base models and calculate the indirect effect

model.xy

```
model.xy <- lm(sleep ~ cond2 + cond3 + anxiety, data=slp_females)
ols_regress(model.xy)
```

model.xm

```
model.xm <- lm(hygiene ~ cond2 + cond3 + anxiety , data=slp_females)
ols_regress(model.xm)
```

model.xmy

```
model.xmy <- lm(sleep ~ cond2 + cond3 + anxiety + hygiene, data=slp_females)
ols_regress(model.xmy)
```

Calculate the indirect effects

```
# Pull out coefficients and standard errors for a and b paths
path.a_cond2 <- coefficients(model.xm)["cond2"]
path.a_cond3 <- coefficients(model.xm)["cond3"]
path.b <- coefficients(model.xmy)["hygiene"]
se.path.a_cond2<-sqrt(vcov(model.xm)["cond2","cond2"])
se.path.a_cond3<-sqrt(vcov(model.xm)["cond3","cond3"])
se.path.b<-sqrt(vcov(model.xmy)["hygiene","hygiene"])

# Compute indirect effects
ab_cond2 <- path.a_cond2*path.b
ab_cond3 <- path.a_cond3*path.b
ab_cond2
ab_cond3
```

Model Results

model.xy

Parameter Estimates

model	Beta	Std. Error	Std. Beta	t	Sig	lower	upper
(Intercept)	78.645	2.118		37.128	0.000	74.472	82.817
cond2	9.182	1.152	0.440	7.969	0.000	6.912	11.451
cond3	13.251	1.090	0.666	12.160	0.000	11.105	15.398
anxiety	-2.657	0.511	-0.257	-5.197	0.000	-3.664	-1.650

Adjusting for anxiety, we see a positive and significant total effect of Condition 2 (compared to Condition 1) and Condition 3 (compared to Condition 1) on sleep efficiency.

model.xm

Parameter Estimates

model	Beta	Std. Error	Std. Beta	t	Sig	lower	upper
(Intercept)	5.644	0.337		16.754	0.000	4.980	6.307
cond2	1.438	0.183	0.464	7.847	0.000	1.077	1.799
cond3	1.801	0.173	0.609	10.393	0.000	1.460	2.143
anxiety	0.011	0.081	0.007	0.134	0.894	-0.149	0.171

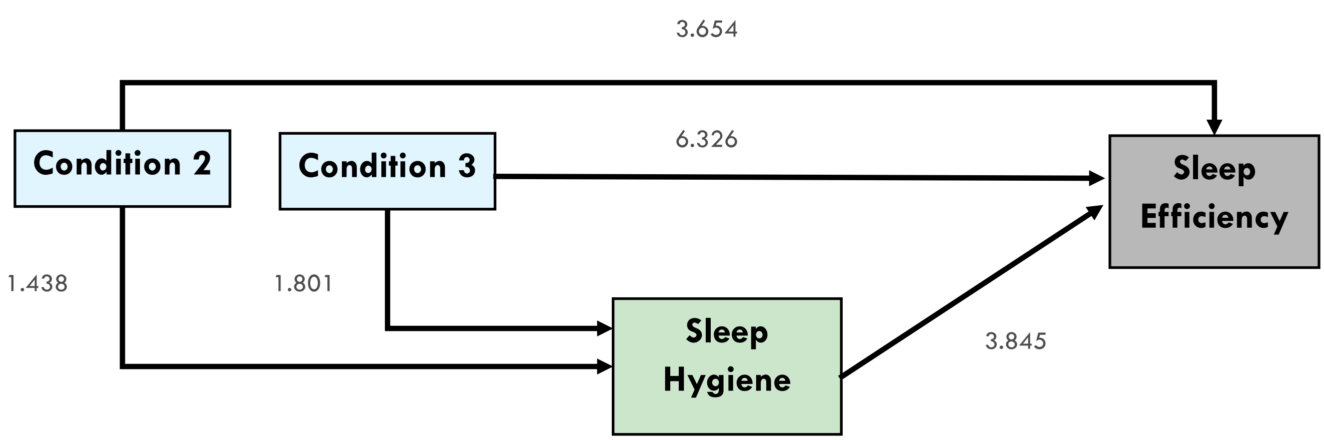
Adjusting for anxiety, we see a positive and significant total effect of Condition 2 (compared to Condition 1) and Condition 3 (compared to Condition 1) on sleep hygiene, indicating two significant a-paths.

model.xmy

Parameter Estimates

model	Beta	Std. Error	Std. Beta	t	Sig	lower	upper
(Intercept)	56.946	2.466		23.096	0.000	52.089	61.803
cond2	3.654	1.023	0.175	3.572	0.000	1.639	5.668
cond3	6.326	1.039	0.318	6.091	0.000	4.280	8.372
anxiety	-2.699	0.405	-0.261	-6.658	0.000	-3.498	-1.901
hygiene	3.845	0.320	0.571	12.021	0.000	3.215	4.475

Adjusting for anxiety and the mediator (hygiene), we see a positive and significant direct effect of Condition 2 (compared to Condition 1) and Condition 3 (compared to Condition 1) on sleep efficiency, though both effects have been substantially reduced compared to model.xy. These are the c'-paths. The beta for hygiene is the b-path, and it is positive and statistically significant, indicating that adjusting for treatment and anxiety, hygiene is associated with sleep efficiency. Since both the a-paths and the b-path are statistically significant, our models meet the minimum requirements for mediation. Now, we can take a look at the significance of the indirect effect.



Indirect effect for Condition 2 = 1.438 • 3.845 = 5.528

Indirect effect for Condition 3 = 1.801 • 3.845 = 6.925

Bootstrap Confidence Intervals

Bootstrap the confidence intervals

```
boot.med <- function(data, indices){
   data <-data[indices,]

 model.xm <- lm(hygiene ~ cond2 + cond3 + anxiety , data=data)
 model.xmy <- lm(sleep ~ cond2 + cond3 + anxiety + hygiene, data=data)

 path.a_cond2 <- coefficients(model.xm)["cond2"]
 path.a_cond3 <- coefficients(model.xm)["cond3"]
 path.b <- coefficients(model.xmy)["hygiene"]
   ab_cond2 <- path.a_cond2*path.b
   ab_cond3 <- path.a_cond3*path.b
   return(c(ab_cond2, ab_cond3))
}

medboot<-boot(data=slp_females, statistic = boot.med, R=10000)
medconfint1 <- boot.ci(medboot, index=1, conf=(.95), type=c("bca"))
medconfint2 <- boot.ci(medboot, index=2, conf=(.95), type=c("bca"))
print (medconfint1)
print (medconfint2)
```

We now have 2 indirect effects, so there are two quantities listed on the return statement in the bootstrap syntax. We bootstrap both quantities (see modconfint1 and medconfint2) in the last chunk of code above. The results are then printed for each quantity specified.

Bootstrap Confidence Intervals Results

```
BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS
Based on 10000 bootstrap replicates

CALL :
boot.ci(boot.out = medboot, conf = (0.95), type = c("bca"), index = 1)

Intervals :
Level       BCa
95%   ( 3.888,  7.435 )
Calculations and Intervals on Original Scale
```

Results for ab_cond2, indirect effect for Condition 2.

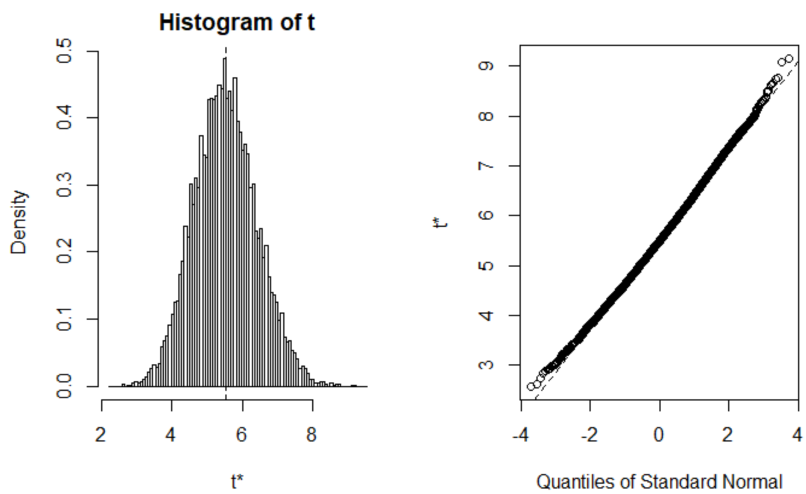
```
BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS
Based on 10000 bootstrap replicates

CALL :
boot.ci(boot.out = medboot, conf = (0.95), type = c("bca"), index = 2)

Intervals :
Level       BCa
95%   ( 5.334,  8.711 )
Calculations and Intervals on Original Scale
```

Results for ab_cond3, indirect effect for Condition 3.

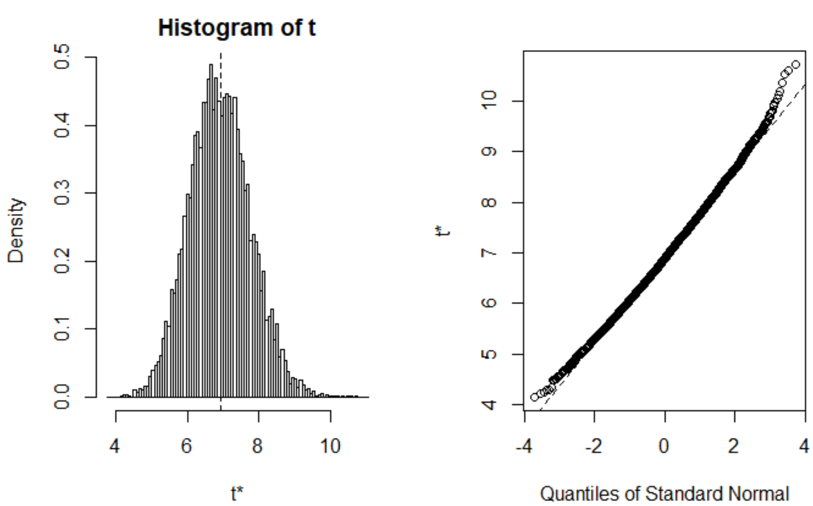
The indirect effect for both conditions is significant (0 is not contained in either confidence interval), meaning that the influence of each program on improved sleep efficiency is, in part, due to sleep hygiene. That is, being in Condition 2 (compared to Condition 1) improved sleep hygiene, which in turn improved sleep efficiency. Likewise, being in Condition 3 (compared to Condition 1) improved sleep hygiene, which in turn improved sleep efficiency. For both conditions, the direct effect is still statistically significant (see model.xmy), indicating the mediation is only partial. This means that part, specifically 5.528/9.182 = .60 (or 60%), of the reason Condition 2 is associated with better sleep efficiency is because of improved sleep hygiene, but a significant direct effect remains. Likewise part, specifically 6.925/13.251 = .52 (or 52%), of the reason Condition 3 is associated with better sleep efficiency is because of improved sleep hygiene, but a significant direct effect remains. Perhaps there are additional mediators to consider (i.e., other mechanisms by which Condition 2 and Condition 3 lead to better sleep efficiency).

Bootstrap Plots for Multiple Returned Quantities

```
plot(medboot, index = 1)
plot(medboot, index = 2)
```



Bootstrap plot for ab_cond2

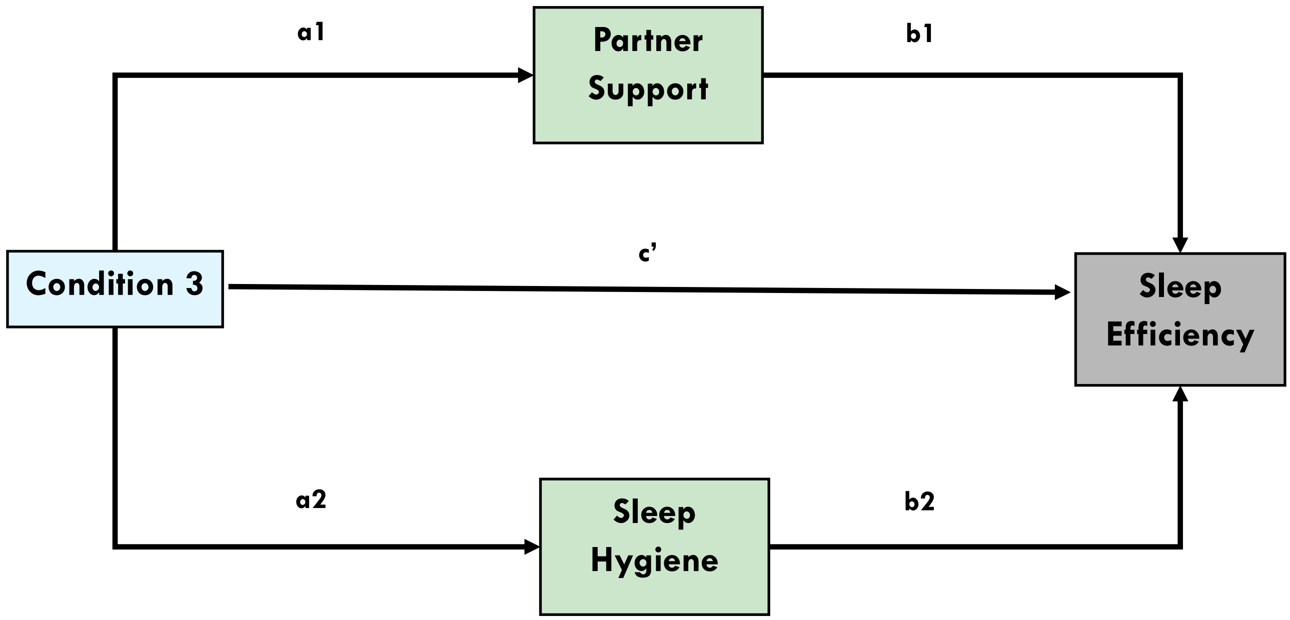


Bootstrap plot for ab_cond3

Advanced Example 2: Mediation With Multiple Mediators

Let's build on the last example. When assessing the indirect effect of Condition 3, we found that hygiene explained part of the effect of the program on sleep efficiency, but not all of it. We concluded that maybe there was an additional mechanism by which the program improved sleep efficiency. Condition 3 included a group-based sleep improvement program that included the participant's partner in the intervention. Perhaps Condition 3 (compared to Condition 1) resulted in better sleep efficiency because the program improved both sleep hygiene practices AND improved the participant's perception that her partner was supportive of her quest to improve sleep. We can simultaneously examine two mediators. For parsimony, we will exclude participants in Condition 2, and consider only females in Conditions 1 and 3.

Now, we have two indirect effects, one through partner support and one through sleep hygiene. When there is more than one mediator, the indirect effect for each unique pathway is called a **specific indirect effect**. We also have a total indirect effect, which is calculated as the sum of the two specific indirect effects. The specific indirect effects describe the unique contribution of each mediator, while the total indirect effect describes the extent to which the effect of Condition 3 on sleep efficiency is explained by both partner support and hygiene.



Specific Indirect Effects:
cond3 → support → efficiency: a1•b1
cond3 → hygiene → efficiency: a2•b2

Total Indirect Effect:
total indirect: a1•b1 + a2•b2

Execute the Models

Subset the data to exclude females in Condition 2

```
slp_females_cond3 <- filter(slp_females, cond2 != 1)
```

Estimate the base models

model.xy

```
model.xy <- lm(sleep ~ cond3, data = slp_females_cond3)
ols_regress(model.xy)
```

model.xm1—for social support

```
model.xm1 <- lm(support ~ cond3, data = slp_females_cond3)
ols_regress(model.xm1)
```

model.xm2—for hygiene

```
model.xm2 <- lm(hygiene ~ cond3, data = slp_females_cond3)
ols_regress(model.xm2)
```

model.xmy

```
model.xmy <- lm(sleep ~ cond3 + support + hygiene, data = slp_females_cond3)
ols_regress(model.xmy)
```

Calculate the indirect effects

```
# Calculate indirect effects
path.a1 <- coefficients(model.xm1)["cond3"]
path.b1 <- coefficients(model.xmy)["support"]
path.a2 <- coefficients(model.xm2)["cond3"]
path.b2 <- coefficients(model.xmy)["hygiene"]

a1b1 <- path.a1*path.b1
a2b2 <- path.a2*path.b2
sumind <- a1b1+a2b2

a1b1
a2b2
sumind
```

Model Results

model.xy

Parameter Estimates

model	Beta	Std. Error	Std. Beta	t	Sig	lower	upper
(Intercept)	68.340	0.790		86.491	0.000	66.781	69.900
cond3	13.021	1.154	0.649	11.285	0.000	10.744	15.298

model.xm1—for support

Parameter Estimates

model	Beta	Std. Error	Std. Beta	t	Sig	lower	upper
(Intercept)	2.882	0.063		45.551	0.000	2.757	3.007
cond3	0.574	0.092	0.425	6.211	0.000	0.392	0.756

model.xm2—for hygiene

Parameter Estimates

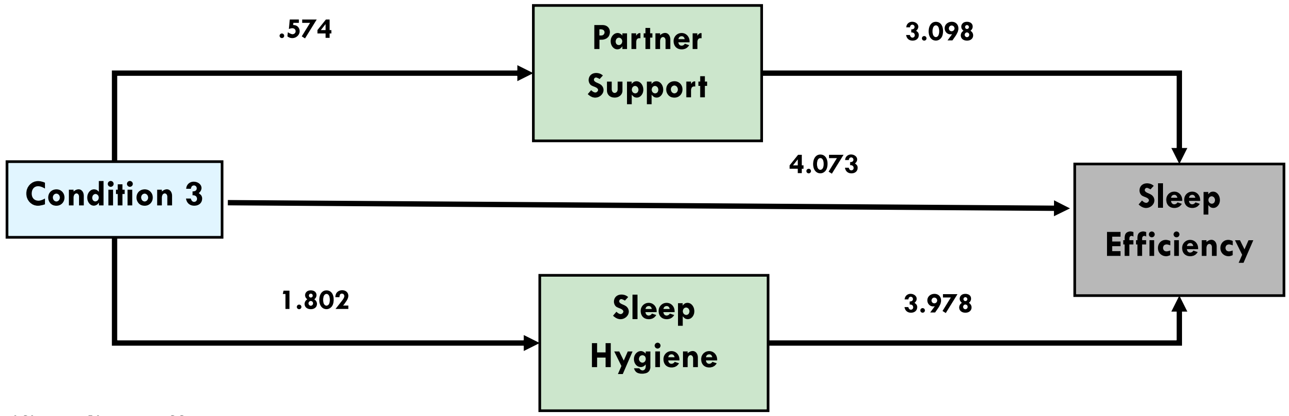
model	Beta	Std. Error	Std. Beta	t	Sig	lower	upper
(Intercept)	5.686	0.115		49.467	0.000	5.459	5.913
cond3	1.802	0.168	0.630	10.737	0.000	1.471	2.134

model.xmy

Parameter Estimates

model	Beta	Std. Error	Std. Beta	t	Sig	lower	upper
(Intercept)	36.792	3.286		11.198	0.000	30.307	43.277
cond3	4.073	1.257	0.203	3.242	0.001	1.593	6.554
support	3.098	0.745	0.208	4.160	0.000	1.628	4.567
hygiene	3.978	0.410	0.567	9.707	0.000	3.169	4.787

Looking at the output for model.xy, we see here that Condition 3 improves both partner support ($b = .574$) and hygiene ($b = 1.802$); therefore both a-paths are statistically significant. In addition, both partner support ($b = 3.098$) and hygiene ($b = 3.978$) are associated with better sleep efficiency, over and above the condition (see model.xmy); therefore, both b-paths are also statistically significant. We have met the initial requirements for mediation for both mediators (i.e, the a- and b-paths are significant for both mediators). Now, we can calculate the product terms for the indirect effects and determine if the indirect effects are statistically significant.



Specific Indirect Effects:

cond3 → support → efficiency: .574 · 3.098 = 1.778

cond3 → hygiene → efficiency: 1.802 · 3.978 = 7.170

Total Indirect Effect:
total indirect: 1.778 + 7.170 = 8.948

Bootstrap Confidence Intervals

Bootstrap the confidence intervals

```
boot.med <- function(data, indices){
  data <-data[indices,]
  model.xy <- lm(sleep ~ cond3, data=data)
  model.xm1 <- lm(support ~ cond3, data=data)
  model.xm2 <- lm(hygiene ~ cond3, data=data)
  model.xmy <- lm(sleep ~ cond3 + support + hygiene, data=data)

  path.a1 <- coefficients(model.xm1)["cond3"]
  path.b1 <- coefficients(model.xmy)["support"]
  path.a2 <- coefficients(model.xm2)["cond3"]
  path.b2 <- coefficients(model.xmy)["hygiene"]

  a1b1 <- path.a1*path.b1
  a2b2 <- path.a2*path.b2
  sumind <- a1b1+a2b2
  return(c(a1b1,a2b2,sumind))
}
medboot<-boot(data = slp_females_cond3, statistic = boot.med, R=10000)
medconfint1 <- boot.ci(medboot, index=1, conf=(.95), type=c("bca"))
medconfint2 <- boot.ci(medboot, index=2, conf=(.95), type=c("bca"))
medconfint3 <- boot.ci(medboot, index=3, conf=(.95), type=c("bca"))
print(medconfint1)
print(medconfint2)
print(medconfint3)
```

```
BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS
Based on 10000 bootstrap replicates

CALL :
boot.ci(boot.out = medboot, conf = (0.95), type = c("bca"), index = 1)

Intervals :
Level       BCa
95%   ( 0.906,  2.929 )
Calculations and Intervals on Original Scale
BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS
Based on 10000 bootstrap replicates

CALL :
boot.ci(boot.out = medboot, conf = (0.95), type = c("bca"), index = 2)

Intervals :
Level       BCa
95%   ( 5.503,  9.260 )
Calculations and Intervals on Original Scale
BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS
Based on 10000 bootstrap replicates

CALL :
boot.ci(boot.out = medboot, conf = (0.95), type = c("bca"), index = 3)

Intervals :
Level       BCa
95%   ( 7.066, 11.222 )
Calculations and Intervals on Original Scale
```

Indirect effect for a1b1

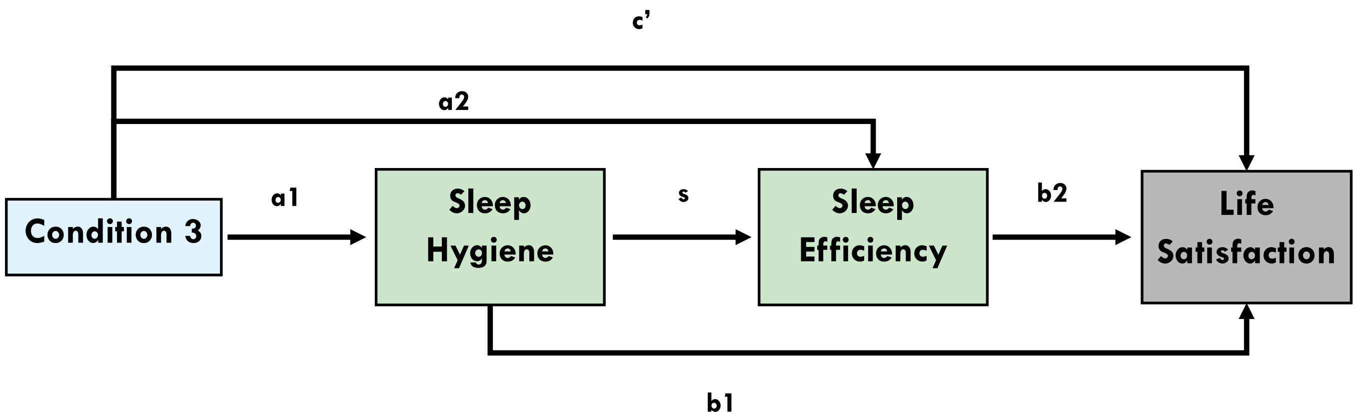
Indirect effect for a2b2

Sum of indirect effects

Both specific indirect effects are statistically significant, as is the total indirect effect. This indicates that both partner support and hygiene are significant unique mediators of the effect of Condition 3 on sleep efficiency. The sum of the indirect effects is 8.948. We can compare this to the total effect (13.021). By dividing the total indirect effect by the total effect we get 8.948/13.021 = .69, so about 69% of the effect of Condition 3 on sleep efficiency is explained by the mediators. We still don't explain all of the variability, as the direct effect in model.xmy for cond3 is still significantly different from 0.

Advanced Example 3: Serial Mediation

We can also examine serial processes of mediation. Here, there are multiple mediators (as was in the case in the last example), but instead of the mediators exerting their influence in parallel, one mediator proceeds another. For example, we might be interested in determining if being in Condition 3 improved sleep hygiene, in turn improved sleep hygiene led to better sleep efficiency, and by enjoying better sleep efficiency, participants reported better life satisfaction. Let's examine this model among females in the study.



Specific Indirect Effects:

cond3 → hygiene → life satisfaction: a1•b1

cond3 → efficiency → life satisfaction: a2•b2

cond3 → hygiene → efficiency → life satisfaction: a1•S•b2

Total Indirect Effect:

total indirect: a1•b1 + a2•b2 + a1•S•b2

Execute the Models

Estimate the base models

model.xy

```
model.xy <- lm(lifesat ~ cond3, data = slp_females_cond3)
ols_regress(model.xy)
```

model.xms1—first section of serial mediation

```
model.xms1 <- lm(hygiene ~ cond3, data = slp_females_cond3)
ols_regress(model.xms1)
```

model.xms2—second section of serial mediation

```
model.xms2 <- lm(sleep ~ cond3 + hygiene, data = slp_females_cond3)
ols_regress(model.xms2)
```

model.xmy

```
model.xmy <- lm(lifesat ~ cond3 + hygiene + sleep, data = slp_females_cond3)
ols_regress(model.xmy)
```

Calculate the indirect effects

```
path.a1 <- coefficients(model.xms1)["cond3"]
path.b1 <- coefficients(model.xmy)["hygiene"]
path.a2 <- coefficients(model.xms2)["cond3"]
path.b2 <- coefficients(model.xmy)["sleep"]
path.s <-  coefficients(model.xms2)["hygiene"]

a1b1 <- path.a1*path.b1
a2b2 <- path.a2*path.b2
a1sb2 <- path.a1*path.s*path.b2
sumind <- a1b1+a2b2+a1sb2

a1b1
a2b2
a1sb2
sumind
```

Model Results

model.xy

Parameter Estimates

model	Beta	Std. Error	Std. Beta	t	Sig	lower	upper
(Intercept)	3.611	0.084		43.248	0.000	3.446	3.776
cond3	0.997	0.122	0.526	8.178	0.000	0.757	1.238

model.xms1

Parameter Estimates

model	Beta	Std. Error	Std. Beta	t	Sig	lower	upper
(Intercept)	5.686	0.115		49.467	0.000	5.459	5.913
cond3	1.802	0.168	0.630	10.737	0.000	1.471	2.134

model.xms2

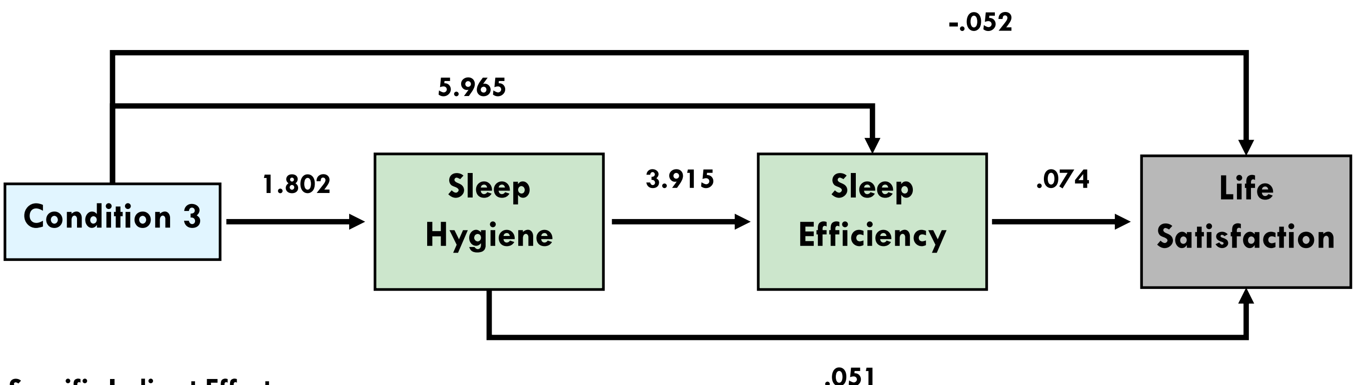
Parameter Estimates

model	Beta	Std. Error	Std. Beta	t	Sig	lower	upper
(Intercept)	46.080	2.521		18.277	0.000	41.104	51.055
cond3	5.965	1.225	0.297	4.870	0.000	3.547	8.383
hygiene	3.915	0.428	0.558	9.140	0.000	3.070	4.760

model.xmy

Parameter Estimates

model	Beta	Std. Error	Std. Beta	t	Sig	lower	upper
(Intercept)	-1.705	0.375		-4.549	0.000	-2.444	-0.965
cond3	-0.052	0.114	-0.028	-0.460	0.646	-0.276	0.172
hygiene	0.051	0.045	0.077	1.123	0.263	-0.039	0.140
sleep	0.074	0.007	0.778	11.153	0.000	0.061	0.087



Specific Indirect Effects:

cond3 → hygiene → life satisfaction: 1.802 · .051 = .092

cond3 → efficiency → life satisfaction: 5.965 · .074 = .439

cond3 → hygiene → efficiency → life satisfaction: 1.802 · 3.915 · .074 = .519

Total Indirect Effect:

total indirect: .092 + .439 + .519 = 1.049

Bootstrap the Confidence Intervals

Bootstrap the confidence intervals

```
boot.med <- function(data, indices){
  data <-data[indices,]
  model.xms1 <- lm(hygiene ~ cond3, data=data)
  model.xms2 <- lm(sleep ~ cond3 + hygiene, data=data)
  model.xmy <- lm(lifesat ~ cond3 + hygiene + sleep, data=data)
  path.a1 <- coefficients(model.xms1)["cond3"]
  path.b1 <- coefficients(model.xmy)["hygiene"]
  path.a2 <- coefficients(model.xms2)["cond3"]
  path.b2 <- coefficients(model.xmy)["sleep"]
  path.s <-  coefficients(model.xms2)["hygiene"]

# Calculate indirect effect
  a1b1 <- path.a1*path.b1
  a2b2 <- path.a2*path.b2
  a1sb2 <- path.a1*path.s*path.b2
  sumind <- a1b1+a2b2+a1sb2
  return(c(a1b1,a2b2,a1sb2,sumind))
}
medboot<-boot(data = slp_females_cond3, statistic = boot.med, R=10000)
medconfint1 <- boot.ci(medboot, index=1, conf=(.95), type=c("bca"))
medconfint2 <- boot.ci(medboot, index=2, conf=(.95), type=c("bca"))
medconfint3 <- boot.ci(medboot, index=3, conf=(.95), type=c("bca"))
medconfint4 <- boot.ci(medboot, index=4, conf=(.95), type=c("bca"))
print(medconfint1)
print(medconfint2)
print(medconfint3)
print(medconfint4)
```

Results of Bootstrap

```
BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS
Based on 10000 bootstrap replicates

CALL :
boot.ci(boot.out = medboot, conf = (0.95), type = c("bca"), index = 1)

Intervals :
Level       BCa
95%   (-0.0552,  0.2521 )
Calculations and Intervals on Original Scale
BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS
Based on 10000 bootstrap replicates

CALL :
boot.ci(boot.out = medboot, conf = (0.95), type = c("bca"), index = 2)

Intervals :
Level       BCa
95%   ( 0.2514,  0.6451 )
Calculations and Intervals on Original Scale
BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS
Based on 10000 bootstrap replicates

CALL :
boot.ci(boot.out = medboot, conf = (0.95), type = c("bca"), index = 3)

Intervals :
Level       BCa
95%   ( 0.3730,  0.6959 )
Calculations and Intervals on Original Scale
BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS
Based on 10000 bootstrap replicates

CALL :
boot.ci(boot.out = medboot, conf = (0.95), type = c("bca"), index = 4)

Intervals :
Level       BCa
95%   ( 0.815,  1.306 )
Calculations and Intervals on Original Scale
```

Indirect effect for a1b1

Indirect effect for a2b2

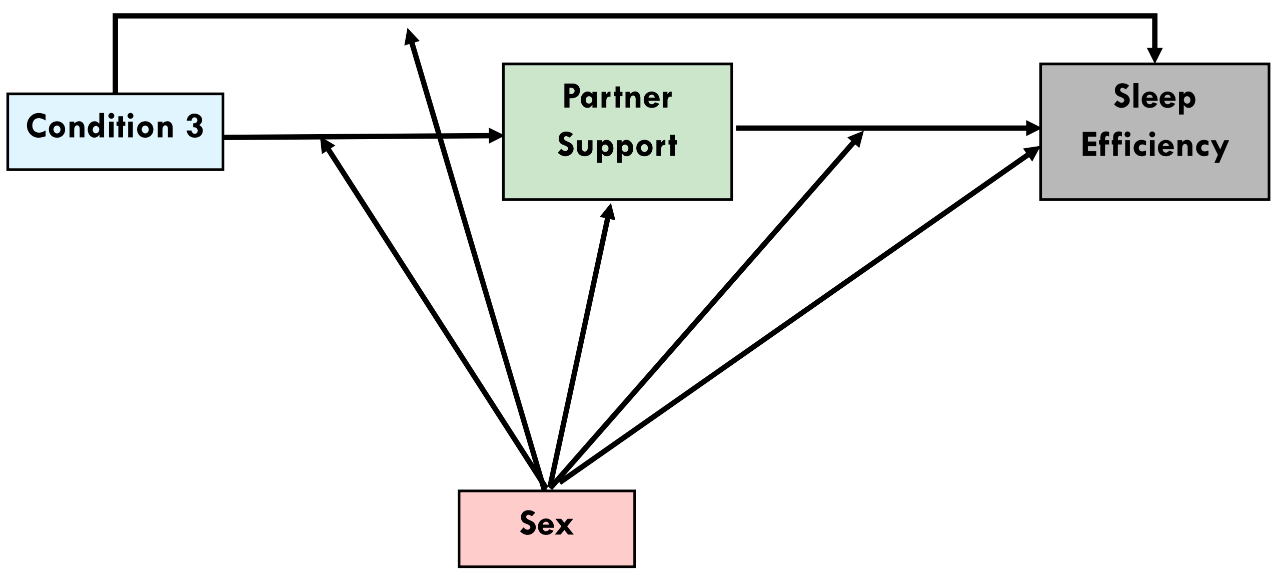
Indirect effect for serial mediation

Sum of indirect effects

All of the effect of Condition 3 on life satisfaction can be explained by these mediators. Notice that no direct effect of Condition 3 on life satisfaction remains in model.xmy. The specific indirect effect of Condition 3 on life satisfaction via hygiene is not statistically significant; however, the specific indirect effect via sleep efficiency is statistically significant and the serial mediation via hygiene and sleep efficiency is statistically significant. In addition the sum of the specific indirect effects is statistically significant.

Advanced Example 4: Moderated Mediation

For our last advanced example, we will consider moderated mediation. Here, one or more of the mediation paths differs as a function of some third variable. Let's determine if the extent to which the indirect effect of Condition 3 on sleep efficiency via partner support is different for males and females.



Execute the Models

Subset the data to keep males and females, but exclude those in Condition 2

```
slp_cond3 <- filter(slp, cond2 != 1)
```

Estimate the base models

model.xy

```
model.xy <- lm(sleep ~ cond3 + female + cond3*female, data = slp_cond3)
ols_regress(model.xy)
linearHypothesis(model.xy, "cond3 + cond3:female")
```

Note that we use the linearHypothesis function from the car package to calculate and test the slopes for females.

model.xm

```
model.xm <- lm(support ~ cond3 + female + cond3*female, data = slp_cond3)
ols_regress(model.xm)
linearHypothesis(model.xm, "cond3 + cond3:female")
```

model.xmy

```
model.xmy <- lm(sleep ~ cond3 + support + female + cond3*female + support*female, data = slp_cond3)
ols_regress(model.xmy)
linearHypothesis(model.xmy, "cond3 + cond3:female")
linearHypothesis(model.xmy, "support + support:female")
```

Calculate the indirect effects

```
path.a_males <- coefficients(model.xm)["cond3"]
path.a_females <- coefficients(model.xm)["cond3"] + coefficients(model.xm)["cond3:female"]
path.b_males <- coefficients(model.xmy)["support"]
path.b_females <- coefficients(model.xmy)["support"] + coefficients(model.xmy)["support:female"]

ab_males <- path.a_males*path.b_males
ab_females <- path.a_females*path.b_females

ab_males
ab_females
```

Model Results

model.xy

Parameter Estimates

model	Beta	Std. Error	Std. Beta	t	Sig	lower	upper
(Intercept)	54.868	0.870		63.038	0.000	53.157	56.579
cond3	17.884	1.202	0.606	14.883	0.000	15.522	20.247
female	13.473	1.270	0.423	10.612	0.000	10.977	15.968
cond3:female	-4.863	1.807	-0.093	-2.691	0.007	-8.416	-1.310

```
Linear hypothesis test

Hypothesis:
cond3  + cond3:female = 0

Model 1: restricted model
Model 2: sleep ~ cond3 + female + cond3 * female

  Res.Df   RSS Df Sum of Sq      F    Pr(>F)
1    397 39274
2    396 31800  1    7473.5 93.065 < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

model.xm

Parameter Estimates

model	Beta	Std. Error	Std. Beta	t	Sig	lower	upper
(Intercept)	2.595	0.049		53.278	0.000	2.499	2.690
cond3	1.242	0.067	0.668	18.470	0.000	1.110	1.374
female	0.287	0.071	-0.033	4.044	0.000	0.148	0.427
cond3:female	-0.668	0.101	-0.234	-6.606	0.000	-0.867	-0.469

```
Linear hypothesis test

Hypothesis:
cond3  + cond3:female = 0

Model 1: restricted model
Model 2: support ~ cond3 + female + cond3 * female

  Res.Df     RSS Df Sum of Sq     F    Pr(>F)
1    397 114.079
2    396  99.562  1    14.517 57.74 2.183e-13 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

model.xmy

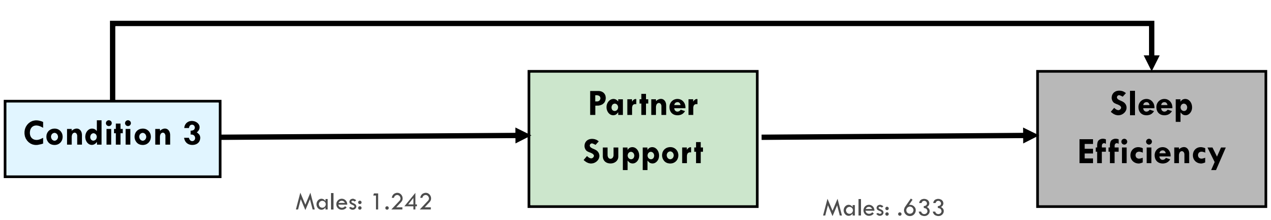
Parameter Estimates

model	Beta	Std. Error	Std. Beta	t	Sig	lower	upper
(Intercept)	53.226	4.074		13.066	0.000	45.217	61.234
cond3	17.098	2.249	0.562	7.604	0.000	12.677	21.519
support	0.633	1.534	0.088	0.413	0.680	-2.383	3.649
female	6.961	5.239	0.427	1.329	0.185	-3.339	17.260
cond3:female	-5.701	2.693	-0.109	-2.117	0.035	-10.996	-0.406
support:female	2.196	1.886	0.059	1.164	0.245	-1.512	5.905

```
Linear hypothesis test

Hypothesis:
cond3  + cond3:female = 0

Model 1: restricted model
Model 2: sleep ~ cond3 + support + female + cond3 * female + support *
    female

  Res.Df   RSS Df Sum of Sq      F    Pr(>F)
1    395 35951
2    394 31260  1    4691.6 59.133 1.189e-13 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
Linear hypothesis test

Hypothesis:
support  + support:female = 0

Model 1: restricted model
Model 2: sleep ~ cond3 + support + female + cond3 * female + support *
    female

  Res.Df   RSS Df Sum of Sq      F  Pr(>F)
1    395 31787
2    394 31260  1    527.16 6.6444 0.01031 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Males: 17.098

Females: 17.098 + -5.701 = 11.397



Males: 1.242

Females: 1.242 + -.668 = .574

Males: .633

Females: .633 + 2.196 = 2.829

Indirect effect for Males = 1.242 • .633 = .787

Indirect effect for Females = .574 • 2.829 = 1.624

The a- and b-paths are both statistically significant for females, indicating that we have met the minimum requirement for mediation. However, while the a-path is significant for males, the b-path is not significant for males. Therefore, partner support is not a mediator for males.

Bootstrap the Confidence Intervals

Bootstrap the confidence intervals

```
boot.med <- function(data, indices){
  data <-data[indices,]
  model.xm <- lm(support ~ cond3 + female + cond3*female, data=data)
  model.xmy <- lm(sleep ~ cond3 + support + female + cond3*female + support*female, data=data)
  path.a_males <- coefficients(model.xm)["cond3"]
  path.a_females <- coefficients(model.xm)["cond3"] + coefficients(model.xm)["cond3:female"]
  path.b_males <- coefficients(model.xmy)["support"]
  path.b_females <- coefficients(model.xmy)["support"] + coefficients(model.xmy)["support:female"]

  # Calculate indirect effect
  ab_males <- path.a_males*path.b_males
  ab_females <- path.a_females*path.b_females
  diff <- ab_males - ab_females
  return(c(ab_males, ab_females, diff))
}
medboot<-boot(data = slp_cond3, statistic = boot.med, R=10000)
medconfint1 <- boot.ci(medboot, index=1, conf=(.95), type=c("bca"))
medconfint2 <- boot.ci(medboot, index=2, conf=(.95), type=c("bca"))
medconfint3 <- boot.ci(medboot, index=3, conf=(.95), type=c("bca"))
print(medconfint1)
print(medconfint2)
print(medconfint3)
```

Mediation is not viable for males given that the b-path is not significant; however, we will calculate the indirect effect for purposes of demonstration. We can also calculate the difference in the indirect effects (see the line that starts with "diff" in the code where the indirect effects are calculated). This is a test of whether the indirect effect is significantly different for males compared to females.

Bootstrap Results

```
BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS
Based on 10000 bootstrap replicates

CALL :
boot.ci(boot.out = medboot, conf = (0.95), type = c("bca"), index = 1)

Intervals :
Level       BCa
95%   (-3.9820,  4.9293 )
Calculations and Intervals on Original Scale
```

Indirect effect for males

```
BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS
Based on 10000 bootstrap replicates

CALL :
boot.ci(boot.out = medboot, conf = (0.95), type = c("bca"), index = 2)

Intervals :
Level       BCa
95%   ( 0.505,   3.082 )
Calculations and Intervals on Original Scale
```

Indirect effect for females

```
BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS
Based on 10000 bootstrap replicates

CALL :
boot.ci(boot.out = medboot, conf = (0.95), type = c("bca"), index = 3)

Intervals :
Level       BCa
95%   (-5.755,   3.554 )
Calculations and Intervals on Original Scale
```

Difference in the indirect effect for males compared to females

The confidence interval for the indirect effect for males contains 0 and therefore is not significant, which is what we would expect given the b-path for males was not significant.

For females the confidence interval for the indirect effect doesn't include 0, indicating that the indirect effect is statistically significant.

The confidence interval for the difference in the indirect effect between males and females is not significant. Therefore, while we have evidence that the indirect effect isn't significant for males, but is significant for females, the test of a differential indirect effect is not significant. We would need a larger sample size to detect this effect.