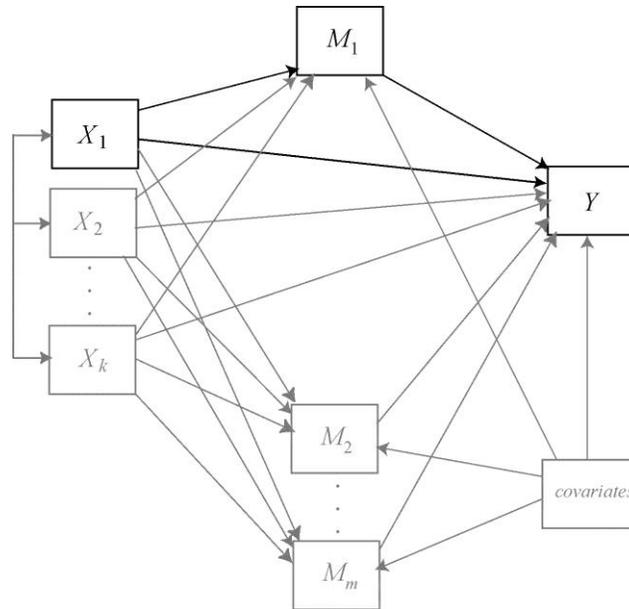


MEDIATE

```
MEDIATE Y = yvar/X = xlist/M = mlist/C = covlist [/SAMPLES = {z} (1000**)]
                                                [/CICONF = {ci} (95**)]
                                                [/TOTAL = {t} (0**)]
                                                [/OMNIBUS = {o} (0**)]
                                                [/CIMETHOD = {m} (1**)]
                                                [/CATX = {cx} {0**}]
                                                [/PERCENT = {p} {0**}]
                                                [/SEED = {sd} {random**}] .
```

Subcommands in brackets are optional
** Default if subcommand is omitted



Overview

MEDIATE estimates the total, direct, and indirect effects of causal variable or variables `xlist` on outcome variable `yvar` through a proposed mediator variable or set of mediator variables `mlist`, controlling for (optional) one or more variables in `covlist`. MEDIATE is similar to INDIRECT (Preacher and Hayes, 2008) but allows multiple X variables and also offers features for handling and coding a single multicategorical X variable. MEDIATE also provides omnibus tests for direct, indirect and total effects for X as set, or the group variable coded with X when is multicategorical. Inferences for indirect effects can be based on either percentile bootstrap confidence intervals or Monte Carlo confidence intervals. Sobel tests (a.k.a. “Normal theory tests”) of indirect effects are not provided as an option in MEDIATE. The principles behind the estimation of direct, indirect, and total effects when X is multicategorical can be found in Hayes and Preacher (2014).

Preparing for Use

The MEDIATE.sps file should be opened as a syntax file in SPSS. Once it has been opened, execute the entire file *exactly as is*. Do not modify the code at all. Once the program is executed, the MEDIATE program window can be closed. You then have access to the MEDIATE command until you quit SPSS. The MEDIATE.sps file must be loaded and reexecuted each time SPSS is opened. See the

“Examples” section below for some examples of how to set up a `MEDIATE` command in a syntax window.

Examples

```
MEDIATE Y=attitude/X=cond/M=commune inter/omnibus=1/samples=5000/catx=3.
```

- Estimates the relative total and relative indirect effects of categorical variable `cond` on `attitude` through `commune` and `inter`.
- Automatically codes the levels of `cond` using sequential coding.
- Generates omnibus tests of direct and indirect effects.
- 95% bootstrap percentile confidence intervals for indirect effects are generated based on 5,000 bootstrap samples.

```
MEDIATE Y=symptoms/X=emotion thought/M=reaction/C=age educ/total=1  
/omnibus=1/ciconf=90/cimethod=2/samples=10000.
```

- Estimates the total, direct, and indirect effects of `emotion` and `thought` on `symptoms` through `reaction` while controlling for `age` and `educ`.
- Prints the model which estimates the total effects.
- Generates 90% Monte Carlo confidence intervals for the indirect effects using 10,000 samples.
- Produces omnibus tests of the total, direct, and indirect effects of `emotion` and `thought`.
- `emotion` and `thought` are treated as either dichotomous or continuous.

Covariates

The direct, indirect, and total effects of `xlist` on `yvar` can be calculated with or without including a set of covariates which are partialled out of `yvar` and all variables in the `mlist` list. The names of covariates should be listed in `covlist`.

Multiple Mediator Variables

`MEDIATE` accepts up to 15 proposed mediator variables assumed to be operating in parallel (i.e., not sequentially linked). Mediators must be continuous or treated as such. `MEDIATE` will automatically detect if any of the mediator variables in `mlist` are dichotomous and, if so, execution will terminate.

Multiple Independent Variables

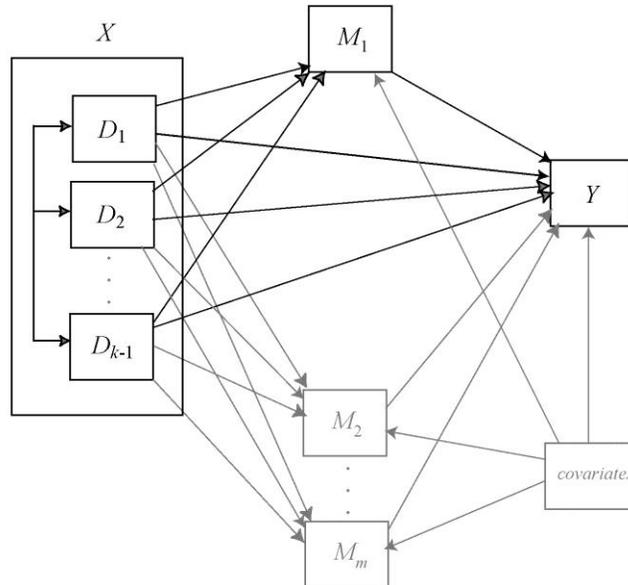
Direct, indirect, and total effects for all variables listed in `xlist` will be estimated unless the `\catx` subcommand is used with an argument (`cx`) greater than 0. These variables are included in the models of the proposed mediators and the outcome simultaneously, so all effects for each variable in `xlist` can be interpreted as independent of or controlling for the other variables in `xlist`. When a `catx` option greater than zero is used, an error message will be displayed and execution will terminate if more than a single variable occurs in `xlist`.

Multicategorical X Variable

Using the `/catx` subcommand, the user can specify an independent variable as categorical. The `/catx` subcommand is available **only when a single variable is listed in `xlist`**. `MEDIATE` assumes when the

`/catx` subcommand is used that the variable listed in `xlist` contains a set of arbitrary numerical codes designating group. A variable specified as categorical can contain up to nine unique codes designating group, and `MEDIATE` will detect the number of distinct groups by counting the number of unique codes. If more than 9 is found, `MEDIATE` will terminate. A dichotomous X need not be specified as categorical, although there is no harm in doing so. A dichotomous variable specified as categorical will be recoded by `MEDIATE` using 0 and 1 as group codes, with the group with the numerical smaller code in the variable in `xlist` coded 0.

When the `/catx` subcommand is used, `MEDIATE` will produce $k - 1$ new variables (D_1 , D_2 , and so forth, in the output) coding membership in one of k mutually exclusive group, with the coding method determined by the argument in `cx`. These $k - 1$ variables represent X in the model and output, as below



In the examples below, assume that the variable in `xlist` is called `cond` and is a categorical variable with four categories coded 1, 3, 4, and 6. When `cx` is set to 1, i.e., `/catx = 1`, simple indicator coding is used with the group coded with the smallest numerical code in `xlist` treated as the reference category. The indicator codes will correspond to groups as coded by the variable in `xlist` in ascending sequential order. For example:

COND	D1	D2	D3
1	0	0	0
3	1	0	0
4	0	1	0
6	0	0	1

When `cx` is set to 2, i.e., `/catx = 2`, effect coding is used, with group with the smallest numerical code left out of the coding scheme. The dummy variables (D_1 , D_2 , and so forth) correspond to groups in ascending sequential order in the coding of the variable in `xlist`. For example:

COND	D1	D2	D3
1	-1	-1	-1
3	1	0	0
4	0	1	0
6	0	0	1

When `cx` is set to 3, i.e., `/catx = 3`, sequential coding is used. Sequential coding is used when the interest is in comparing the mean of group j to the mean of group $j + 1$. `MEDIATE` will assume that the ascending ordinality of the independent variable corresponds with the ascending sequence of arbitrary numerical codes in the variable in `xlist`. For example:

COND	D1	D2	D3
1	0	0	0
3	1	0	0
4	1	1	0
6	1	1	1

When `cx` is set to 4, i.e., `/catx = 4`, Helmert coding is used, with the groups coded such that they differ by one unit. Helmert coding is useful when interest is in comparing group j to the mean of all groups ordinally higher than group j . `MEDIATE` will assume that the ascending ordinality of the independent variable corresponds with the ascending sequence of arbitrary numerical codes in the variable in `xlist`. For example:

COND	D1	D2	D3
1	-0.75	0	0
3	0.25	-0.667	0
4	0.25	0.333	-0.5
6	0.25	0.333	0.5

Helmert coding is also useful for setting up certain orthogonal contrasts for a nominal categorical independent variable.

There is no option in `MEDIATE` for changing which group is treated as the reference. If you want to designate a different group as the reference group, recode the variable in `xlist` prior to using `MEDIATE` so that the reference group you desire is coded with the numerically smallest code.

Inference for Indirect Effects

`MEDIATE` produces either percentile bootstrap, bias-corrected bootstrap, or Monte Carlo confidence intervals for inference about indirect effects using the `\samples` subcommand. An indirect effect can be interpreted as different from zero with confidence `ci%` if zero is outside of the confidence interval. The default is bias-corrected percentile bootstrap confidence intervals. Percentile bootstrap confidence intervals can be requested by setting the `p` argument in the `\PERCENT` subcommand to 1 (i.e., `/PERCENT = 1`). Monte Carlo confidence intervals are produced by setting the `m` argument in the `CIMETHOD` subcommand to 2 (i.e., `/CIMETHOD = 2`). By default, confidence level is set at 95. This can be changed by setting the `ci` argument in the `/CICONF` subcommand to any desired confidence between 50 and up to but not including 100 (e.g., `/CICONF = 90`). Noninteger confidence levels can be specified. Only one type of confidence interval can be constructed at once.

Bootstrap and Monte Carlo confidence intervals are both based on random sampling of either the data itself (with replacement, in the case of bootstrapping) or from normal distributions with means and standard errors defined by the point estimates and standard errors of the paths that define the indirect effects (in the case of Monte Carlo intervals). The number of samples is set to 1000 by default, but more

is recommended. The number of samples can be changed through the `z` argument in the `/SAMPLES` subcommand (i.e., `/SAMPLES = 5000`).

When the `/catx` subcommand is used to specify the independent variable as categorical, the user has the option of requesting stratified bootstrapping for construction of confidence intervals for indirect effects by setting the `m` argument to 3, i.e., `/CIMETHOD = 3`. With stratified bootstrapping, the bootstrap samples are generated by stratifying the resampling in each group coded in the variable in `xlist`. For instance, if the categorical variable has three categories with frequencies 20, 30, and 40, then each bootstrap sample will always contain 20 cases sampled with replacement from the first group, 30 from the second, and 40 from the third group.

It is possible to seed the random number generator in order to reproduce the same set of random bootstrap or Monte Carlo samples from run to run of `MEDIATE`. By default, the random number generator is seeded randomly. To specify a seed, using any integer between 1 and 2,000,000,000 as the argument for `sd` in the `/SEED` subcommand (e.g., `/SEED = 3423`).

The point estimates of all indirect effects are displayed by `MEDIATE`, along with the upper and lower bounds of the confidence interval. In addition, a standard error for the indirect effect is produced as the standard deviation of the `z` bootstrap or Monte Carlo estimates.

So called “Normal theory” tests of indirect effects, such as the Sobel test, are not available as an option in `MEDIATE`. See Hayes (2009) and Preacher and Hayes (2004, 2008) for a discussion of the problems with normal theory tests.

“Total Effects” Model

Using the `/TOTAL` subcommand and setting `t` to 1 (i.e., `/TOTAL = 1`) tells `MEDIATE` to print the total effects model in the output. The total effects model is a model of the outcome variable (`yvar`) when estimated from just the variables in `xlist` and `covlist`. The variable(s) in `mlist` are excluded from this model.

Omnibus Tests

`MEDIATE` will produce various omnibus tests of total, direct, and indirect effects using the `/OMNIBUS` subcommand, setting the `t` argument to 1 (i.e., `/OMNIBUS = 1`). An omnibus test is used to answer the question as whether there is evidence that variable or variable(s) X exerts an effect on Y without specifying which variable in the set of X variables is responsible for the effect or, in the case of a multicategorical X , the nature of the difference between group means that is responsible for that effect.

An omnibus test of the direct effect of X is conducted by ascertaining whether the addition of the independent variable(s) in `xlist` to a model of `yvar` containing only proposed mediators in `mlist` and covariates in `covlist` improves the fit of the model, as indexed by a change in the squared multiple correlation that results when the `xlist` variables are added. The increase in R^2 is transformed to a statistic distributed as $F(k, df_2)$ under the null hypothesis of no direct effect, where k is the number of X variables in the model and df_2 is the residual degrees of freedom from the larger model that includes the k variables in `xlist`. When X is a multicategorical variable coding k groups, of interest is whether the inclusion of the $k - 1$ variables coding group improves the fit of the model. The change in R^2 is transformed to a statistic distributed as $F(k - 1, df_2)$ under the null hypothesis of no direct effect. This

test is equivalent to a test of mean group differences in analysis of covariance, controlling for the covariates (if any) and the proposed mediators.

The omnibus test for indirect effect through mediator M_i is based on the index

$$\varphi = (R_{M_i, X}^{\hat{\phi}})(b_i)$$

where $R_{M_i, X}^{\hat{\phi}}$ is adjusted R^2 from the model estimating mediator variable M_i from the k variables in `xlist` or, in the case of a categorical X , the $k - 1$ variables coding group. When covariates are in the model, $R_{M_i, X}^{\hat{\phi}}$ is the change in adjusted R^2 when the X variables are added. b_i is the regression coefficient estimating `yvar` from M_i controlling for `covlist` and all variables in `xlist`. The index does not have a substantive interpretation, but it has an expected value very close to zero when there is no partial association between X and M and/or between M and Y . For inference, `MEDIATE` will produce a bootstrap confidence interval for φ . If the confidence interval does not contain zero, then one can claim an indirect effect of X (as a set if more than one X is specified) on Y through M_i .

An omnibus test for the total effect is produced only when `/OMNIBUS` is used in conjunction with `/TOTAL = 1`. This test asks whether the inclusion of the variables in `xlist` improves the estimation of `yvar` when added to a model containing only the covariates listed in `covlist`. Under the null hypothesis of no total effect of the X variable(s), the increase in R^2 , after transformation to an F ratio, follows the $F(k, df_2)$ distribution, where k is the number of variables in `xlist` and df_2 is the residual degrees of freedom for the model containing the `xlist` and `covlist` variables. If no covariates are listed, the omnibus test is equivalent to the test of the overall total effects model as displayed in the model summary. When X is a multicategorical variable, the omnibus total effect test answers the question as whether there is a difference between the k groups on Y on average independent of (if specified) any covariates in the model. It is equivalent to a test of mean group differences in analysis of covariance, controlling for the covariates. The change in R^2 after transformation follows the $F(k - 1, df_2)$ distribution under the null hypothesis.

Homogeneity of Regression

`MEDIATE` automatically conducts a test of the assumption of homogeneity of the effect of M on Y across values of X , also known as the assumption of *homogeneity of regression* in analysis of covariance or the “no interaction” assumption in mediation analysis. Rejection of this assumption implies that X and M interact in the model of Y . When more than one mediator is specified in the `/m= list`, `MEDIATE` produces an omnibus test of the null hypothesis that X does not interact with *any* of the mediators as well as a test that X and M_i interact assuming no interaction between X and any other mediator. The direct and indirect effects of X on Y should not be interpreted if there is evidence of interaction between X and a mediator listed in the `m= list`.

Categorical Mediator or Dependent Variable

`MEDIATE` should not be used to estimate effects in models with a binary or multinomial outcome (Y) variable or mediator (M).

Notes

- The variables in `mvlist` and `yvar` must be quantitative variables and are assumed to have at least interval-level measurement properties. Variables in `xlist` and `covlist` can be dichotomous or quantitative with interval-level properties. Using the `catx` option, `xlist` can be multicategorical.
- No more than 15 mediators can be included in `mvlist`.
- When using the `catx` option, the multicategorical variable listed as *X* in `xlist` can have no more than 9 discrete values.
- A case will be deleted from the analysis if missing on any of the variables in the model.
- Do not use STRING formatted variables in any of your models. Doing so will produce errors. All variables should be NUMERIC format.
- All coefficients in the output are unstandardized and estimated using ordinary least squares regression.

References

Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York: The Guilford Press. <http://www.guilford.com/p/hayes3>

Hayes, A. F. (2009). Beyond Baron and Kenny: Statistical mediation analysis in the new millennium. *Communication Monographs*, *76*, 408-420.

Hayes, A. F., & Preacher, K. J. (2014). Statistical mediation analysis with a multicategorical independent variable. *British Journal of Mathematical and Statistical Psychology*, *67*, 451-470.

Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in multiple mediator models. *Behavior Research Methods, Instruments, and Computers*, *37*, 717-731.

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, *40*, 879-891.