# INCORPORATING META-SCIENCE IN THE DEVELOPMENT OF MEDIATION AND MODERATION ANALYSIS

AMANDA KAY MONTOYA
DEPARTMENT OF PSYCHOLOGY
UNIVERSITY OF CALIFORNIA, LOS ANGELES

Slides available at github.com/akmontoya/USC2023

# TODAY'S TALK

**Two-Instance Repeated-Measures Designs**

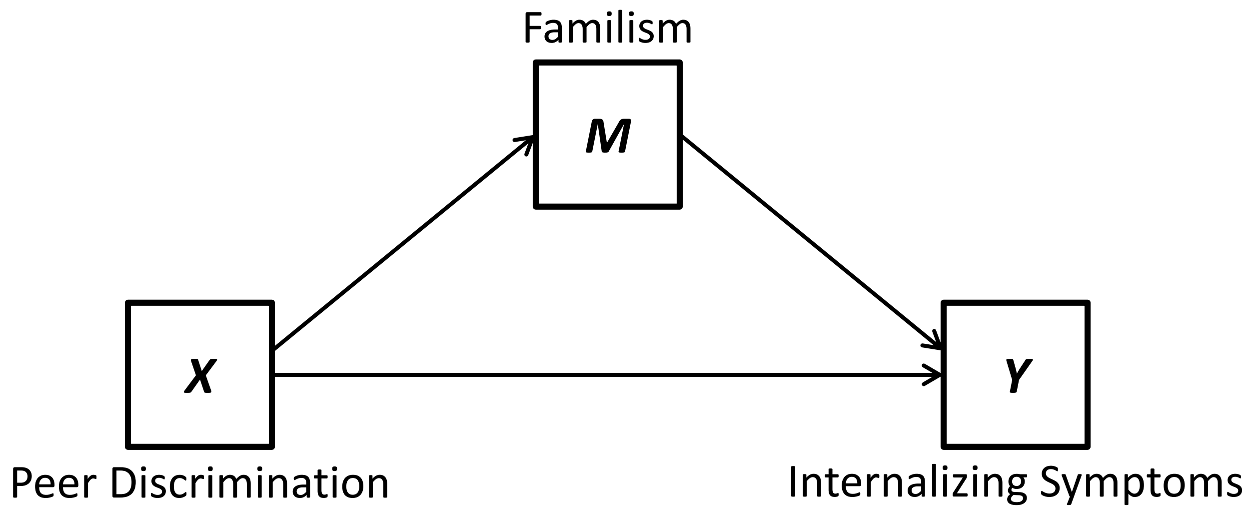
- Mediation Analysis
- Moderation Analysis
- Conditional Process Analysis

**Meta-Science**

- Quality of Evidence in Mediation Analysis
- Sample Size Planning in Mediation and Moderated Mediation
- Registered Reports

**A Call to Arms**

# MEDIATION: WHAT EXPLAINS AN EFFECT?



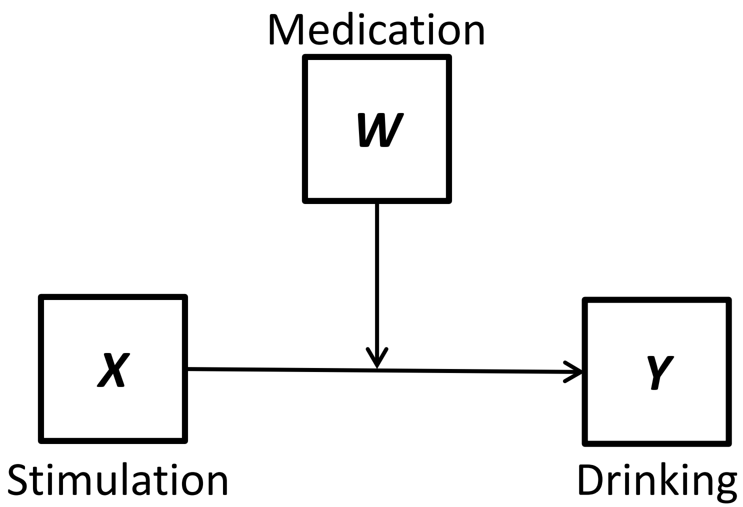

A simple mediation model connects an **assumed** causal variable (*X*) to an **assumed** outcome variable (*Y*), through some mechanism (*M*).

**Example:** Among a sample of LatinX youth, we examined if peer discrimination (*X*) affects internalizing symptoms (*Y*) through a reduction in familism (*M*).

The goal of statistical mediation analysis is to determine if there is an effect of *X* on *Y* through *M*. This is typically done by estimating the ***indirect effect*** and testing if it is different than zero.

Ramos, Delgadillo, Fossum, Montoya, Thamrin, Rapp, Escovar, Chavira (2021)

# MODERATION: WHEN AND FOR WHOM?




The relationship between the focal predictor (*X*) and an outcome (*Y*) is said to be moderated when the size or direction of this relationship depends on *W*.

**Example**: Can medication (ibudilast vs. placebo, *W*) alter the effect of stimulation during drinking (*X*) on future drinking behavior (*Y*)?

Moderation helps us understand ***boundary conditions*** of effects: for whom or when is the effect large or small, present or absent, positive or negative.

Meredith, Grodin, Montoya, Miranda, Squeglia, Towns, Evans, Ray, 2022

# META-SCIENCE: THE SCIENCE OF SCIENCE

Meta-science is the **scientific study of the practice of science**

Increased in popularity about the **replicability crisis** as a way to reflect on how common practices in our field impact the fidelity of our results.

**Meta-scientific methods:**

- Large-scale systematic literature reviews
- Monte Carlo Simulation
- (Multi-site) Replication
- Survey Methods
- Experimental Designs
- Qualitative methods (e.g., focus groups)

My work blends meta-science methods with quantitative methods to **reflect on research practices** while simultaneously **developing new research practices**.

# TWO-INSTANCE REPEATED-MEASURES DESIGNS

# TWO-INSTANCE REPEATED-MEASURES DESIGNS

The causal variable of interest is the factor which differs by repeated measures.

*X*: varies between repeated measurements
*M*: measured in each of the two instances
*Y*: measured in each of the two instances

**Examples:**

- Participants read two scenarios. Interested in how scenario influences *Y* through *M*. Measure *M* and *Y* in each scenario.
- Pre-post test: A group of participants will all go through same intervention, measure hypothesized mediator and outcome before and after treatment.

# FACTORS AFFECTING WOMEN'S INTEREST IN STEM

Developing a strong STEM workforce is a top priority in the US

- NSF INCLUDES: "Transforming education and career pathways to help broaden participation in science and engineering"

**Stereotypes** of STEM fields can lead women to be opt out

**Cultural ideologies** communicate ways to select an occupation

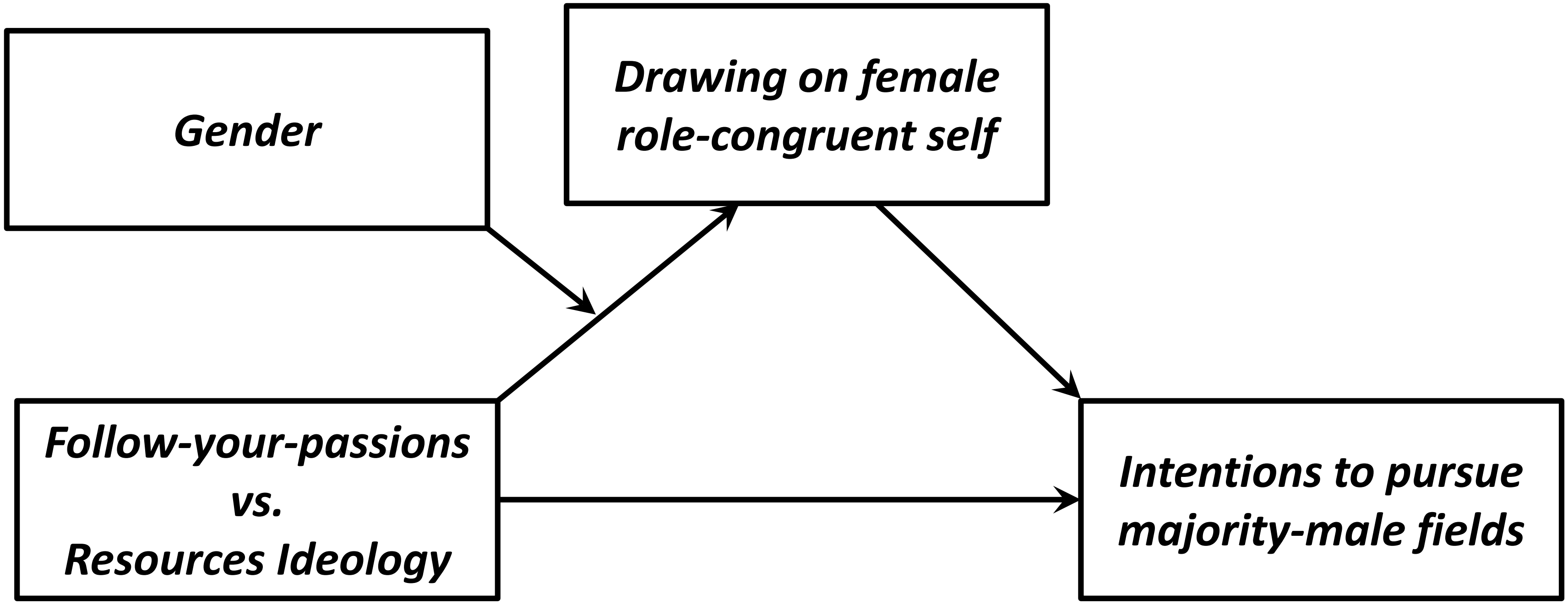
- "Follow your passions" is a common ideology in the US
- "Resource driven" ideology is common in places where individualism is less prominent
  - These places also have greater proportions of women in STEM

**Our Question:** Why would the follow-your-passions ideology decrease women's interest in STEM?

Siy, Germano, Vianna, Azpeitia, Yan, Montoya, Cheryan, in press, *JPSP*

# NEGATIVE IMPACTS OF "FOLLOWING YOUR PASSIONS"

Our theoretical model




Siy, Germano, Vianna, Azpeitia, Yan, <u>Montoya,</u> Cheryan, in press, *JPSP*

# THE DATA

**A two-instance repeated-measures design was used**

Participants responded in each of the two conditions (order counterbalanced)

*People are sometimes told to pursue a career that…*

*allows you to follow your passions*

OR

*leads to a high income*

"List a career that would fit this ideology"

**Outcome:** 3 questions on 1 (women) – 7 (men) Likert scale, averaged

"To what extent is the career you listed typically associated with females or males in U.S. society?"

**Mediator:** 3 questions on 1 (masculine) – 7 (feminine) Likert scale, averaged

"To what extent does the advice above cause you to draw on aspects of yourself that are feminine or masculine (regardless of your gender)?"

Siy, Germano, Vianna, Azpeitia, Yan, Montoya, Cheryan, in press, *JPSP*

# MONTOYA & HAYES, 2017


Psychological Methods
2017, Vol. 22, No. 1, 6–27

1082-989X/17/$12.00 http://dx.doi.org/10.1037/met0000086

## Two-Condition Within-Participant Statistical Mediation Analysis: A Path-Analytic Framework

Amanda K. Montoya and Andrew F. Hayes
The Ohio State University

Researchers interested in testing mediation often use designs where participants are measured on a dependent variable *Y* and a mediator *M* in both of 2 different circumstances. The dominant approach to assessing mediation in such a design, proposed by Judd, Kenny, and McClelland (2001), relies on a series of hypothesis tests about components of the mediation model and is not based on an estimate of or formal inference about the indirect effect. In this article we recast Judd et al.'s approach in the path-analytic framework that is now commonly used in between-participant mediation analysis. By so doing, it is apparent how to estimate the indirect effect of a within-participant manipulation on some outcome through a mediator as the product of paths of influence. This path-analytic approach eliminates the need for discrete hypothesis tests about components of the model to support a claim of mediation, as Judd et al.'s method requires, because it relies only on an inference about the product of paths—the indirect effect. We generalize methods of inference for the indirect effect widely used in between-participant designs to this within-participant version of mediation analysis, including bootstrap confidence intervals and Monte Carlo confidence intervals. Using this path-analytic approach, we extend the method to models with multiple mediators operating in parallel and serially and discuss the comparison of indirect effects in these more complex models. We offer macros and code for SPSS, SAS, and Mplus that conduct these analyses.

*Keywords:* mediation, indirect effect, path analysis, within-participant design, resampling methods



Statistical mediation analysis allows an investigator to answer questions about the process by which some presumed causal variable *X* operates to affect an outcome variable *Y*. Using simple principles of linear modeling (though other analytical approaches are possible; Imai, Keele, & Tingley, 2010; Pearl, 2010, 2012), mediation analysis is used to quantify and test the pathways of influence from *X* to *Y*. In a mediation process, one of those pathways consists of a sequence of causal steps in which *X* affects a *mediator* variable *M*, which in turn causally influences *Y*. This *indirect effect of X*—the conjunction of the effect of *X* on *M* and the effect of *M* on *Y*—quantifies the degree to which *M* acts as the "mechanism" by which *X* affects *Y*. An indirect effect that is different from zero by an inferential test is used to support (but by no means definitively establishes or proves) a claim of mediation of *X*'s effect on *Y* by *M*.

Mediation analysis is commonplace in the social sciences, business, medical research, and many other areas. For example, White, Abu-Rayya, Bliuc, and Faulkner (2015) investigated how long-term interaction with a member of the same religion or a different religion (*X*) influenced intergroup bias (*Y*) through five different emotions (e.g., anger and sadness; *M*). Littleton (2015) found that pregnant women who had a history of sexual victimization (*X*) had higher rates of depression (*M*), which predicted increased somatic complaints (e.g., back pain; *Y*). Schuldt, Guillory, and Gay (2016) examined how the weight of a person recommending a recipe (*X*) influenced the perceived healthiness of the recipe (*Y*) through the perceived health of the recommender (*M*).

Discussions of mediation analysis and its application are most typically couched in terms of or conducted using data from research designs that are cross-sectional or "between-participant" in nature. Typically in these designs, participants are measured once on a proposed mediator *M* and dependent variable *Y*, as in the examples above. This may occur following random assignment of participants into one of two conditions (*X*) that vary via some manipulation (e.g., a "treatment" vs. a "control" group) that is presumed to cause differences in *M* and *Y*. Alternatively, measurement of *M* and *Y* may occur contemporaneously with the observation of *X* (rather than random assignment). For expositional convenience, we refer to designs of this sort (i.e., with or without random assignment to *X*) throughout this article as "between-participant" designs.

Less attention in the methodology literature has been dedicated to mediation analysis when the data come from repeated measurement of the same people on variables in the mediation process, even though such designs are common. In this article we address mediation analysis in a specific category of repeated measures designs. Researchers sometimes measure a dependent variable *Y* and a mediator *M* in two different situations or circumstances (*X*),

This article was published Online First June 30, 2016.
Amanda K. Montoya and Andrew F. Hayes, Department of Psychology, The Ohio State University.
We extend our appreciation to Dr. Simone Dohle and Dr. Michael Siegrist for their generosity in providing access to their data and permission for us to reproduce it in this article. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant DGE-1343012.
Correspondence concerning this article should be addressed to Amanda K. Montoya or Andrew F. Hayes, Department of Psychology, The Ohio State University, 1827 Neil Ave Mall, Columbus, OH 43210. E-mail: montoya.29@osu.edu or hayes.338@osu.edu



- Path-analytic approach to mediation in two-instance repeated-measures designs
- Criticisms of stepwise approach (Judd, Kenny, and McClelland, 2001)
- Generalization to multiple mediator models
  - Parallel Mediation
  - Serial Mediation
- Introduces MEMORE, and SPSS and SAS macro for estimation and inference in these models



Montoya & Hayes, 2017

# PATH ANALYTIC METHOD FOR MEDIATION

Total Effect $c$:

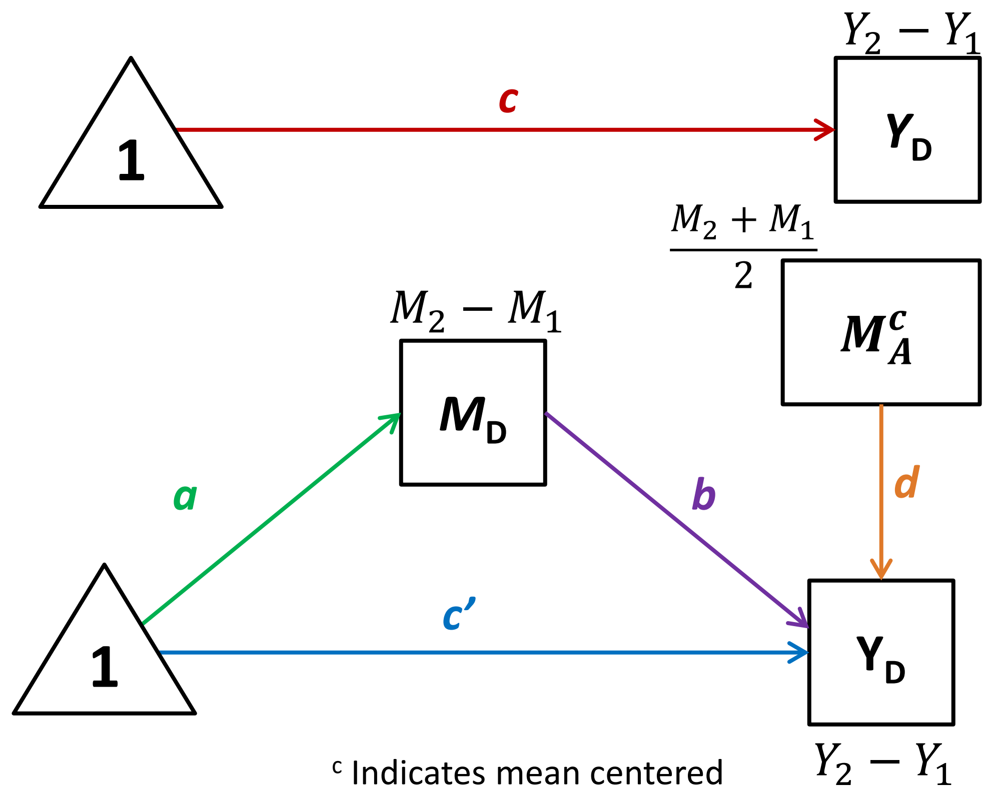
$$Y_{2i} - Y_{1i} = c + \epsilon_{Y_i^*}$$

$a$ path:

$$M_{2i} - M_{1i} = a + \epsilon_{M_i}$$

$b$ path and $c'$ path:

$$Y_{Di} = c' + bM_{Di} + dM_{Ai}^{c} + \epsilon_{Y_i}$$




Indirect effect of *instance* on $Y$ (through $M$) = $a \times b$

Direct effect of *instance* on $Y$ (not through $M$) = $c'$

Total effect = direct effect + indirect effect

$$c = c' + a \times b$$

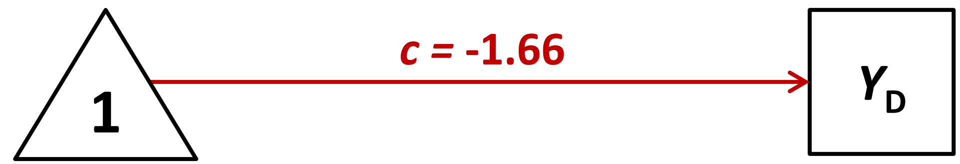
Note: $M_A$ must be mean centered for $c'$ to have intended interpretation

Montoya & Hayes, 2017

# FOLLOW YOUR PASSIONS (WOMEN ONLY)

Total Effect *c*:

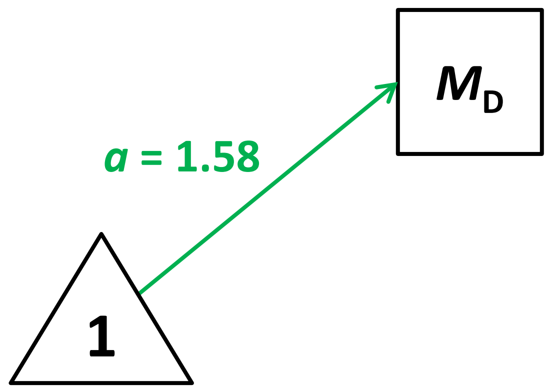
$$Y_{2i} - Y_{1i} = -1.66 + \epsilon_{Y_i^*}$$




Women rated careers generated under the follow-your-passions ideology as **1.66** units less associated with men than careers generated under the resources ideology ($t(374) = 18.80$, $p < .00001$).

*a* path:

$$M_{2i} - M_{1i} = 1.58 + \epsilon_{M_i}$$




Women drew on their female-role congruent selves **1.58** units more when considering the follow-your-passions ideology than when considering the resources ideology ($t(374) = 18.21$, $p < .00001$)



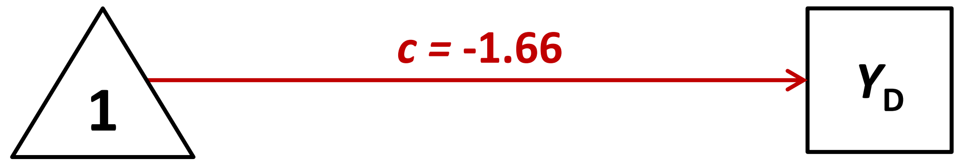
Montoya & Hayes, 2017

Siy, Germano, Vianna, Azpeitia, Yan, Montoya, Cheryan, in press, *JPSP*

# FOLLOW YOUR PASSIONS (WOMEN ONLY)

Total Effect *c*:

$$Y_{2i} - Y_{1i} = -1.66 + \epsilon_{Y_i^*}$$

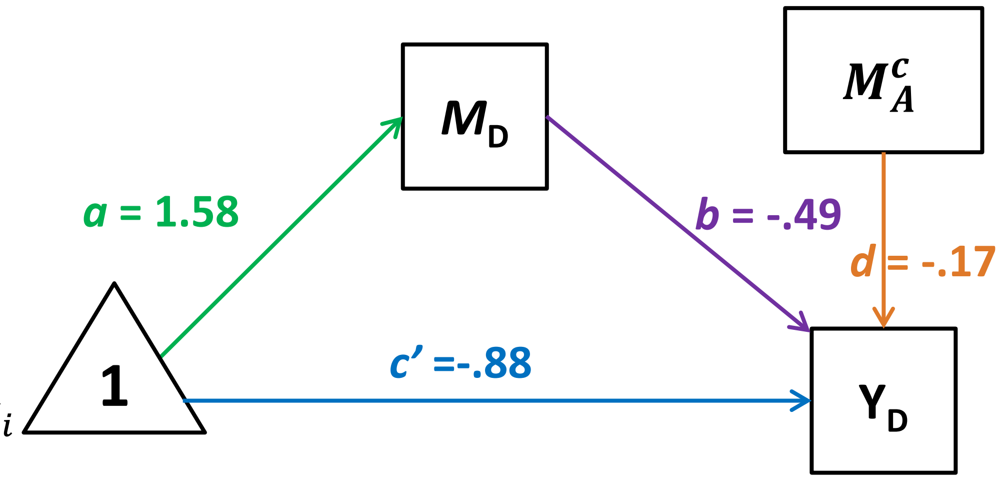



*a* path:

$$M_{2i} - M_{1i} = 1.58 + \epsilon_{M_i}$$

*b* path, *d* path, and *c'* path:

$$Y_{Di} = -.88 - .49M_{Di} - .17M_{Ai}^{c} + \epsilon_{Y_i}$$




$^{c}$ Indicates mean centered

Independent of the effect of drawing on your female role congruent self, careers generated in the resources condition are expected to be .88 units less feminine than those generated in the passions condition ($t(372) = -8.29$, $p < .0001$).



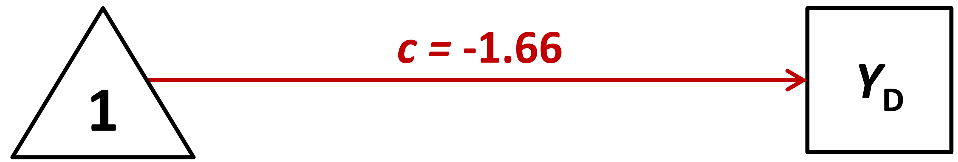
Montoya & Hayes, 2017

Siy, Germano, Vianna, Azpeitia, Yan, Montoya, Cheryan, in press, *JPSP*

# FOLLOW YOUR PASSIONS (WOMEN ONLY)

Total Effect *c*:

$$Y_{2i} - Y_{1i} = -1.66 + \epsilon_{Y_i^*}$$

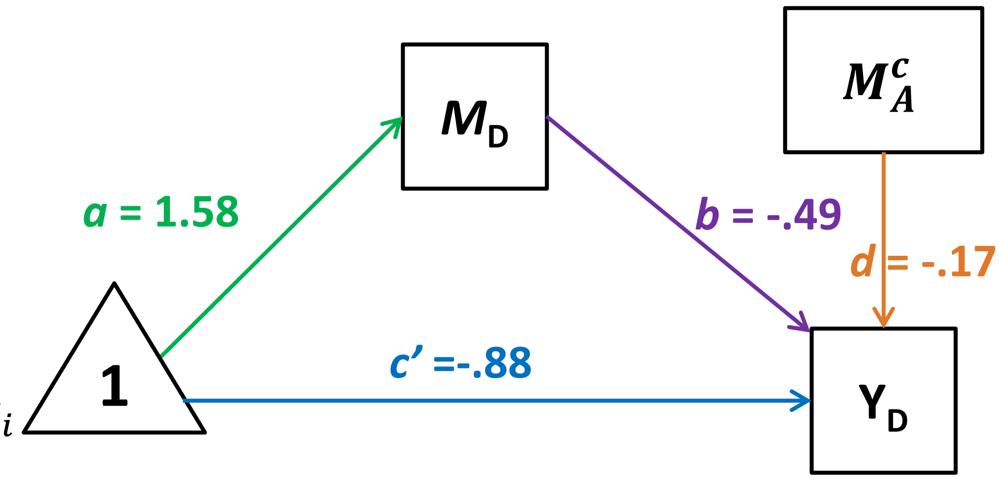



*a* path:

$$M_{2i} - M_{1i} = 1.58 + \epsilon_{M_i}$$

*b* path, *d* path, and *c'* path:

$$Y_{Di} = -.88 - .49 M_{Di} - .17 M_{Ai}^{c} + \epsilon_{Y_i}$$




$^{c}$ Indicates mean centered

Women who draw on their female role-congruent selves more rated the generated careers as less feminine ($b = -.49$, $t(372) = -10.63$, $p < .0001$)

The negative impact of drawing on their female-role congruent self on interest is stronger (more negative) in the follow-your-passions condition than the resource driven condition ($d = -.17$, $t(372) = -1.72$, $p = .09$)



Montoya & Hayes, 2017

Siy, Germano, Vianna, Azpeitia, Yan, Montoya, Cheryan, in press, *JPSP*

# INTERPRETTING THE COEFFICIENTS

In this model there are two measures of the effect of *M* on *Y:* $g_{11}$ and $g_{21}$

$$\boldsymbol{E}(Y_{1i}) = g_{10} + g_{11}M_{1i}$$
$$\boldsymbol{E}(Y_{2i}) = g_{20} + g_{21}M_{2i}$$

Subtract these equations, to get the effect of the difference in *M* on the difference in *Y*.

$$\boldsymbol{E}(Y_{1i} - Y_{2i}) = g_{10} - g_{20} + g_{11}M_{1i} - g_{21}M_{2i}$$

Then apply a rotation to get:

$$\boldsymbol{E}(Y_{1i} - Y_{2i}) = g_{10} - g_{20} + \frac{g_{21} + g_{11}}{2}(M_{1i} - M_{2i}) + (g_{11} - g_{21})\frac{M_{1i} + M_{2i}}{2}$$

$b = \dfrac{g_{21} + g_{11}}{2}$ — *b* is the relationship between *M* and *Y* averaged across the two conditions

$d = (g_{11} - g_{21})$ — *d* is the difference in the relationship between *M* and *Y* between the two conditions (*XM* interaction)



Montoya & Hayes, 2017

# THE INDIRECT EFFECT

Total Effect *c*:

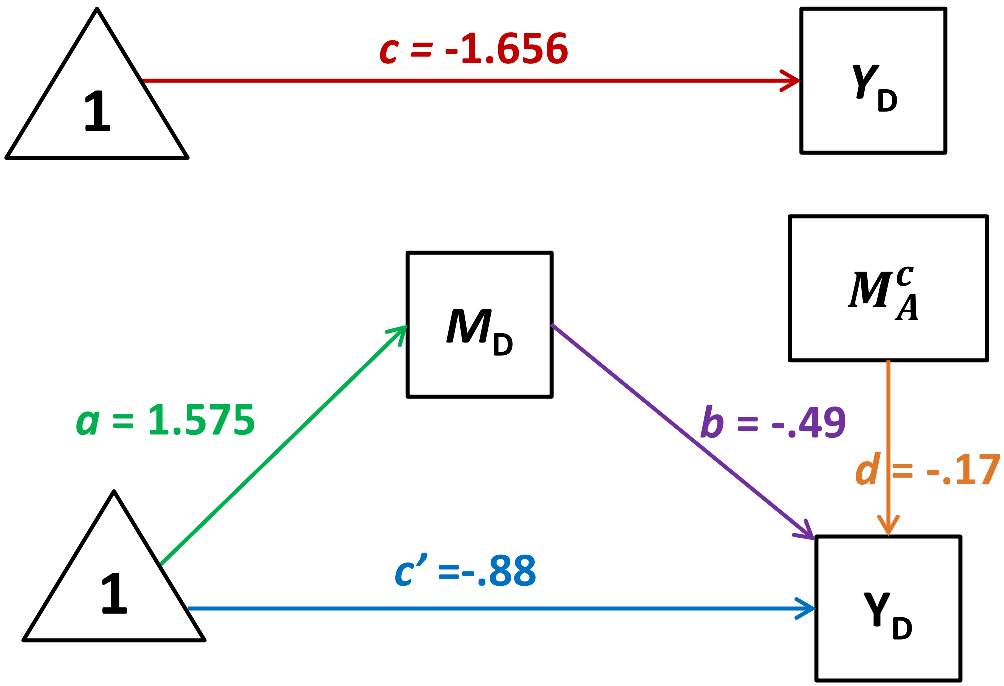
$$Y_{2i} - Y_{1i} = -1.656 + \epsilon_{Y^*i}$$

*a* path:

$$M_{2i} - M_{1i} = 1.575 + \epsilon_{Mi}$$

*b* path, *d* path, and *c'* path:

$$Y_D = -.88 - .49 M_{\mathrm{D}} - .17 M_{\mathrm{A}}^{c} + \epsilon_{Yi}$$




[c] Indicates mean centered

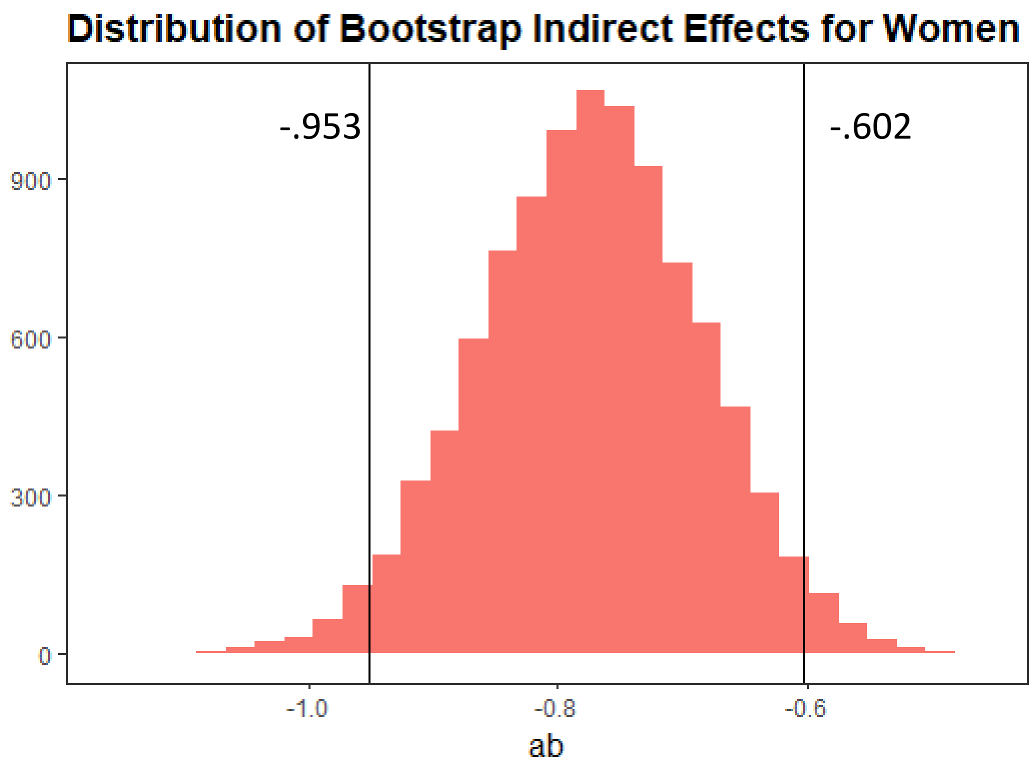
Indirect effect of *instance* on *Y* (through *M*) = $1.575 \times -.49 = -.77$

Women expected careers in the follow-your-passions condition to be .77 less associated with men compared to the resources driven condition, due to drawing on female-role congruent selves more in the passions condition which in turn decreased expected masculinity of careers.

Montoya & Hayes, 2017

Siy, Germano, Vianna, Azpeitia, Yan, Montoya, Cheryan, in press, *JPSP*

# INFERENCE ABOUT THE INDIRECT EFFECT




Zero is not contained in the confidence interval [-.953, -.602] so we conclude the indirect effect is different from zero with 95% confidence.

Among women, there is a significant indirect effect of the follow-your-passions ideology on intentions to pursue male dominated careers through drawing on the female role-congruent self.

Siy, Germano, Vianna, Azpeitia, Yan, Montoya, Cheryan, in press, *JPSP*

# MEMORE

MEMORE is a macro available for SPSS and SAS for conducting (<u>ME</u>diation and <u>MO</u>deration in <u>RE</u>peated measures designs). Documentation and download at akmontoya.com. Mediation functions described in *Two-Condition Within-Participant Statistical Mediation Analysis: A Path-Analytic Framework*

Psychological Methods
2017, Vol. 22, No. 1, 6–27


1082-989X/17/$12.00 http://dx.doi.org/10.1037/met0000086

## Two-Condition Within-Participant Statistical Mediation Analysis: A Path-Analytic Framework

Amanda K. Montoya and Andrew F. Hayes
The Ohio State University

Researchers interested in testing mediation often use designs where participants are measured on a dependent variable *Y* and a mediator *M* in both of 2 different circumstances. The dominant approach to assessing mediation in such a design, proposed by Judd, Kenny, and McClelland (2001), relies on a series of hypothesis tests about components of the mediation model and is not based on an estimate of or formal inference about the indirect effect. In this article we recast Judd et al.'s approach in the path-analytic framework that is now commonly used in between-participant mediation analysis. By so doing, it is apparent how to estimate the indirect effect of a within-participant manipulation on some outcome through a mediator as the product of paths of influence. This path-analytic approach eliminates the need for discrete hypothesis tests about components of the model to support a claim of mediation, as Judd et al.'s method requires, because it relies only on an inference about the product of paths—the indirect effect. We generalize methods of inference for the indirect effect widely used in between-participant designs to this within-participant version of mediation analysis, including bootstrap confidence intervals and Monte Carlo confidence intervals. Using this path-analytic approach, we extend the method to models with multiple mediators operating in parallel and serially and discuss the comparison of indirect effects in these more complex models. We offer macros and code for SPSS, SAS, and Mplus that conduct these analyses.

*Keywords:* mediation, indirect effect, path analysis, within-participant design, resampling methods



Statistical mediation analysis allows an investigator to answer questions about the process by which some presumed causal variable *X* operates to affect an outcome variable *Y*. Using simple principles of linear modeling (though other analytical approaches are possible; Imai, Keele, & Tingley, 2010; Pearl, 2010, 2012), mediation analysis is used to quantify and test the pathways of influence from *X* to *Y*. In a mediation process, one of those pathways consists of a sequence of causal steps in which *X* affects a *mediator* variable *M*, which in turn causally influences *Y*. This *indirect effect of X*—the conjunction of the effect of *X* on *M* and the effect of *M* on *Y*—quantifies the degree to which *M* acts as the "mechanism" by which *X* affects *Y*. An indirect effect that is different from zero by an inferential test is used to support (but by no means definitively establishes or proves) a claim of mediation of *X*'s effect on *Y* by *M*.

Mediation analysis is commonplace in the social sciences, business, medical research, and many other areas. For example, White, Abu-Rayya, Bliuc, and Faulkner (2015) investigated how long-term interaction with a member of the same religion or a different religion (*X*) influenced intergroup bias (*Y*) through five different emotions (e.g., anger and sadness; *M*). Littleton (2015) found that pregnant women who had a history of sexual victimization (*X*) had higher rates of depression (*M*), which predicted increased somatic complaints (e.g., back pain; *Y*). Schuldt, Guillory, and Gay (2016) examined how the weight of a person recommending a recipe (*X*) influenced the perceived healthiness of the recipe (*Y*) through the perceived health of the recommender (*M*).

Discussions of mediation analysis and its application are most typically couched in terms of or conducted using data from research designs that are cross-sectional or "between-participant" in nature. Typically in these designs, participants are measured once on a proposed mediator *M* and dependent variable *Y*, as in the examples above. This may occur following random assignment of participants into one of two conditions (*X*) that vary via some manipulation (e.g., a "treatment" vs. a "control" group) that is presumed to cause differences in *M* and *Y*. Alternatively, measurement of *M* and *Y* may occur contemporaneously with the observation of *X* (rather than random assignment). For expositional convenience, we refer to designs of this sort (i.e., with or without random assignment to *X*) throughout this article as "between-participant" designs.

Less attention in the methodology literature has been dedicated to mediation analysis when the data come from repeated measurement of the same people on variables in the mediation process, even though such designs are common. In this article we address mediation analysis in a specific category of repeated measures designs. Researchers sometimes measure a dependent variable *Y* and a mediator *M* in two different situations or circumstances (*X*),

This article was published Online First June 30, 2016.

Amanda K. Montoya and Andrew F. Hayes, Department of Psychology, The Ohio State University.

We extend our appreciation to Dr. Simone Dohle and Dr. Michael Siegrist for their generosity in providing access to their data and permission for us to reproduce it in this article. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant DGE-1343012.

Correspondence concerning this article should be addressed to Amanda K. Montoya or Andrew F. Hayes, Department of Psychology, The Ohio State University, 1827 Neil Ave Mall, Columbus, OH 43210. E-mail: montoya.29@osu.edu or hayes.338@osu.edu

6

MEMORE can estimate a variety of models. Find the model you are interested in in the templates file, then use that model number.

SPSS Syntax:

```
MEMORE Y = Y1 Y2 /M = M1 M2 /model = 1.
```

SAS Syntax:

```
MEMORE(data = filename, Y = Y1 Y2, M = M1 M2, model = 1);
```



Montoya & Hayes, 2017

```
********************** MEMORE Procedure for SPSS Version 3.0 **********************
                        Written by Amanda Montoya
                  Documentation available at akmontoya.com
***************************************************************************************
Model:
  1
Variables:
Y = FMpas    FMinc
M = gendSpas gendSinc
Computed Variables:
Ydiff =          FMpas     -       FMinc
Mdiff =          gendSpas  -       gendSinc
Mavg =  (        gendSpas  +       gendSinc )        /2                        Centered
Sample Size:
  375
Seed:
  453078
***************************************************************************************
Outcome: Ydiff =  FMpas     -       FMinc
Model
             Effect          SE           t           p        LLCI        ULCI
constant   -1.65600      .08810   -18.79632      .00000    -1.82924    -1.48276
Degrees of freedom for all regression coefficient estimates:
  374
***************************************************************************************
Outcome: Mdiff =  gendSpas  -       gendSinc
Model
             Effect          SE           t           p        LLCI        ULCI
constant    1.57511      .08649    18.21107      .00000     1.40504     1.74518
Degrees of freedom for all regression coefficient estimates:
  374
***************************************************************************************
Outcome: Ydiff =  FMpas     -       FMinc
Model Summary
          R       R-sq        MSE           F         df1         df2           p
     .48415     .23440    2.24046    56.94638     2.00000   372.00000      .00000
Model
              coeff          SE           t           p        LLCI        ULCI
constant    -.88079      .10629    -8.28668      .00000    -1.08979     -.67178
Mdiff       -.49216      .04632   -10.62539      .00000     -.58324     -.40108
Mavg        -.17329      .10058    -1.72286      .08574     -.37107      .02449
Degrees of freedom for all regression coefficient estimates:
  372
*********************** TOTAL, DIRECT, AND INDIRECT EFFECTS ***********************
Total effect of X on Y
     Effect         SE          t         df          p       LLCI       ULCI
   -1.65600     .08810  -18.79632  374.00000     .00000   -1.82924   -1.48276
Direct effect of X on Y
     Effect         SE          t         df          p       LLCI       ULCI
    -.88079     .10629   -8.28668  372.00000     .00000   -1.08979    -.67178
Indirect Effect of X on Y through M
          Effect     BootSE   BootLLCI   BootULCI
Ind1     -.77521     .08928    -.95259    -.60240
Indirect Key
Ind1  'X'      ->       Mdiff     ->       Ydiff
```

Model Information

Total Effect Model

Model for $M_2$-$M_1$

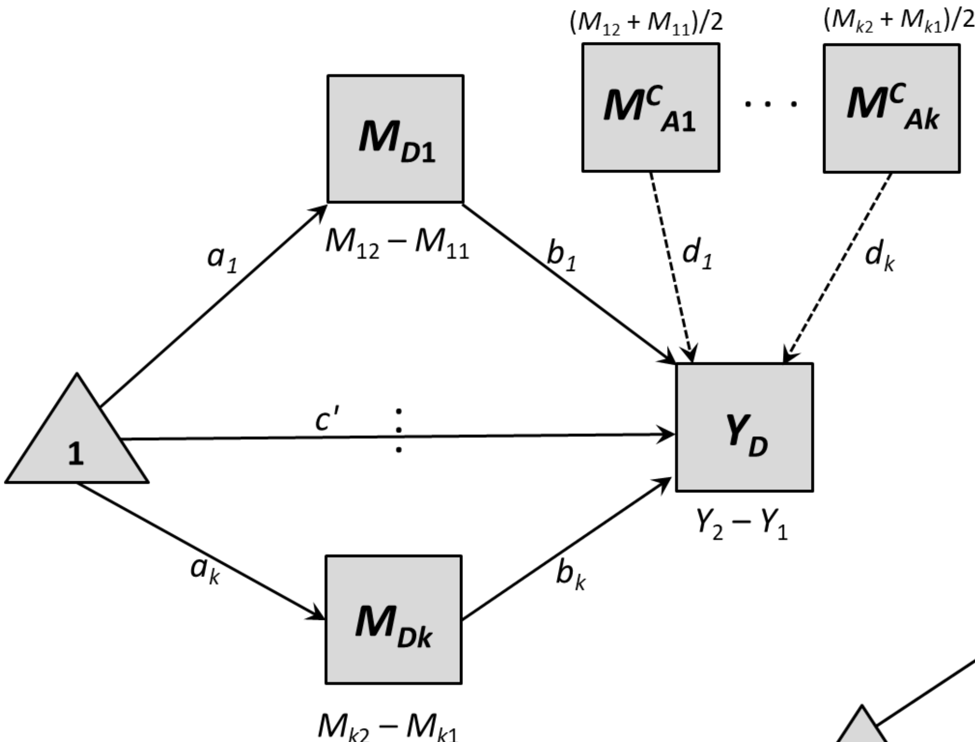
Model for $Y_2$-$Y_1$

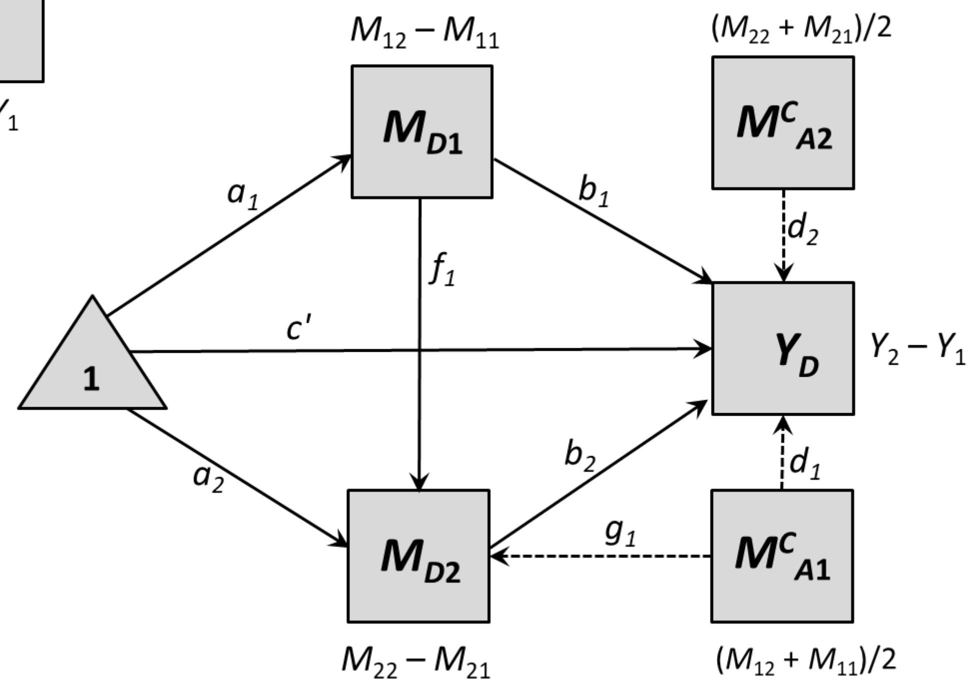
Total, Direct, and Indirect Effects



Montoya & Hayes, 2017

# MULTIPLE MEDIATOR MODELS: PARALLEL & SERIAL



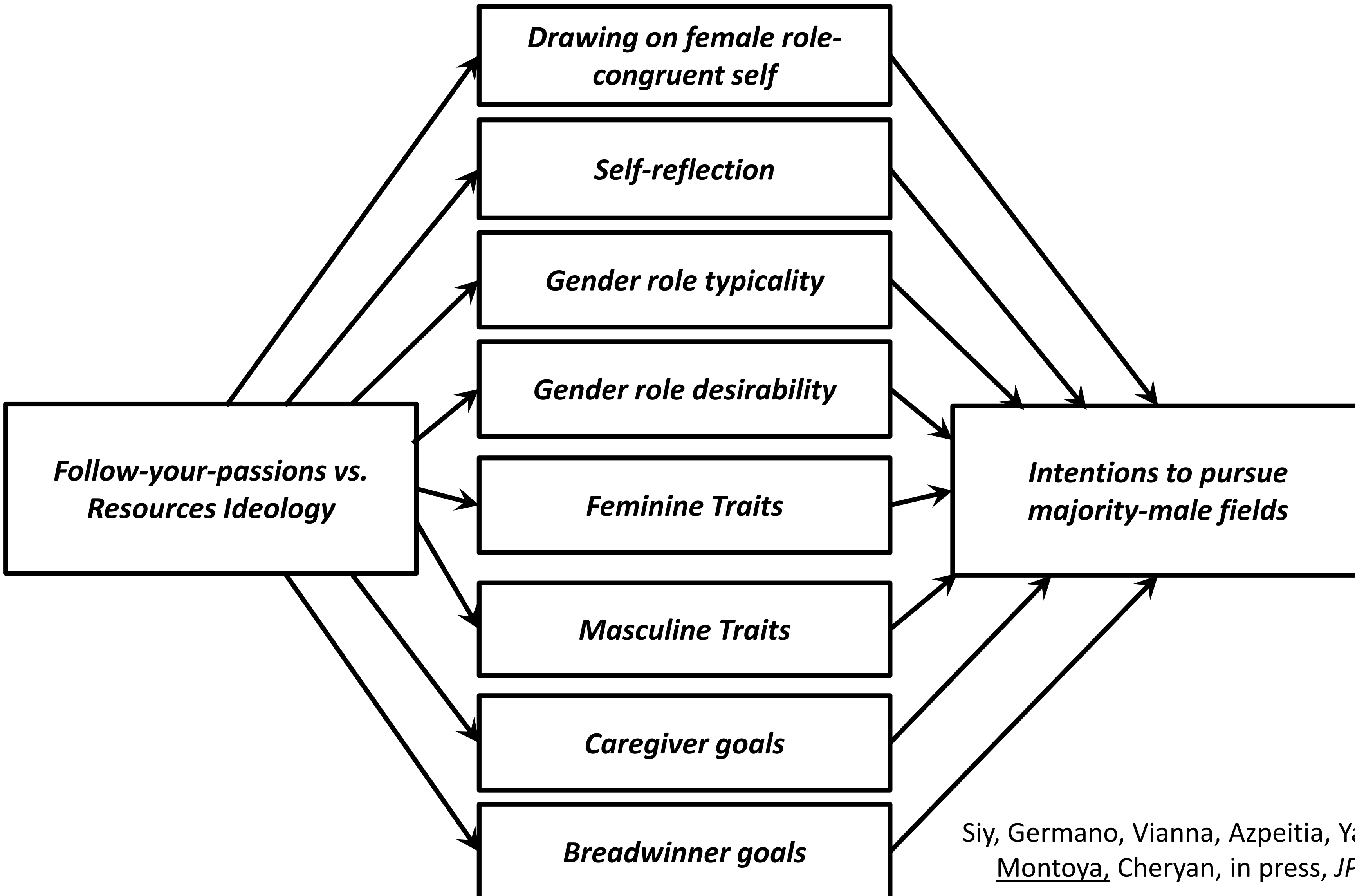






Montoya & Hayes, 2017

# FOLLOW YOUR PASSIONS

**We tested parallel mediators in the follow-your-passions research**




Siy, Germano, Vianna, Azpeitia, Yan, Montoya, Cheryan, in press, *JPSP*

# ADVANCES

Model based approach, rather than piecewise hypotheses

Provides an estimate of the indirect, total, and direct effects

- Allows us to conduct inferential tests directly on an estimate of the indirect effect
- Montoya (2022) Introduced an R script for **power analysis** in these models

Connects researchers understanding of between-subjects mediation to within-subjects mediation

- Reduce misinterpretation of regression coefficients

Using a path analytic framework will help extend to more complicated questions

- Multiple mediators
- Moderated mediation
- Integration of between and within-subjects designs

Montoya, 2022

Montoya & Hayes, 2017

# SOME EXTENSIONS (IN THE WORKS)

**Dichotomous outcomes (Nickie Yang & Jessica Fossum)**

What if, at the end of the study, we asked participants to **choose which career** they preferred?

- Derived estimates of indirect effects
- Compared results from a study with continuous and dichotomous outcomes
- Software implementation
- Tutorial

**More than two conditions (Alondra Cruz)**

In one study, we also looked at the *communal ideology*.

- Derived estimates of *relative indirect effects* (3 conditions)
- Applied Example
- General solution for any contrasts
- Software implementation

# MODERATION IN TWO INSTANCE REPEATED MEASURES DESIGNS

Judd, C. M., McClelland, G. H., and Smith, E. R. (1996). Testing Treatment by Covariate Interactions When Treatment Varies Within Subjects. *Psychological Methods*, *1*(4), 366-378.

Psychological Methods
1996, Vol. 1, No. 4, 366–378



## Testing Treatment by Covariate Interactions When Treatment Varies Within Subjects

Charles M. Judd and Gary H. McClelland
University of Colorado at Boulder

Eliot R. Smith
Purdue University

In contrast to the situation when an independent or treatment variable varies between subjects, procedures for testing treatment by covariate interactions are not commonly understood when the treatment varies within subjects. The purpose of this article is to identify analytic approaches that test such interactions. Two design scenarios are discussed, one in which the covariate is measured only a single time for each subject and hence varies only between subjects, and the other in which the covariate is measured at each level of the treatment variable and hence varies both within and between subjects. In each case, alternative analyses are identified and their assumptions and relative efficiencies compared.

An issue that arises with some frequency in data analysis in psychological research concerns the relationship between some measured variable and the dependent variable and whether that relationship depends on or varies across levels of a manipulated or experimental independent variable. For instance, in a clinical intervention study, we might randomly assign patients to one of two conditions, either a treatment intervention or a placebo intervention control condition. Prior to treatment, we measure a characteristic of the patients, probably focusing on the prior course and severity of their illness. Following the treatment, we assess the outcome variable of symptom severity. The primary question of interest, of course, is whether the outcome variable is affected by the manipulated treatment: Did the treatment make a difference on subsequent symptom severity? Additionally, however, we may well want to know whether the relationship between the treatment and posttreatment symptom severity depends on the patient's pretreatment course of illness. It may be, for instance, that the treatment's effect is greater for patients whose pretreatment symptoms were relatively severe. Equivalently, it may be that posttreatment symptom severity is less well predicted by pretreatment course of illness in the case of patients in the intervention condition than in the case of patients in the control condition.

The pretreatment measure of illness course is typically called a *covariate*. The analysis that is of interest is an analysis of covariance (ANCOVA), including the treatment by covariate interaction (Judd & McClelland, 1989). The two questions of interest are (a) Is there an overall treatment main effect? and (b) Is there a Treatment × Covariate interaction? If the interaction is significant, it indicates that the covariate:outcome variable relationship depends on the treatment variable. Equivalently, it suggests that the effect of the treatment on the outcome variable depends on the level of the covariate.

The analysis is readily conducted using multiple regression, making the standard assumption that errors or residuals are independently sampled from a single normally distributed population. Assume that $Y_i$ is the outcome variable, $Z_i$ is the covariate, and $X_i$ is the contrast-coded (Judd & McClelland, 1989; Rosenthal & Rosnow, 1985) treatment variable. One estimates two least squares regression models:

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 Z_i + \varepsilon_i$$

and

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 Z_i + \beta_3 X_i Z_i + \varepsilon_i.$$

In the first equation, $\beta_1$ represents the magnitude of

Charles M. Judd and Gary H. McClelland, Department of Psychology, University of Colorado at Boulder; Eliot R. Smith, Department of Psychological Sciences, Purdue University.

This work was partially supported by National Institute of Mental Health Grant R01 MH45049.

Correspondence concerning this article should be addressed to Charles M. Judd, Department of Psychology, University of Colorado, Boulder, Colorado 80309. Electronic mail may be sent via the Internet to charles.judd@colorado.edu.





**Does the degree to which *W* predicts *Y* depend on instance?**

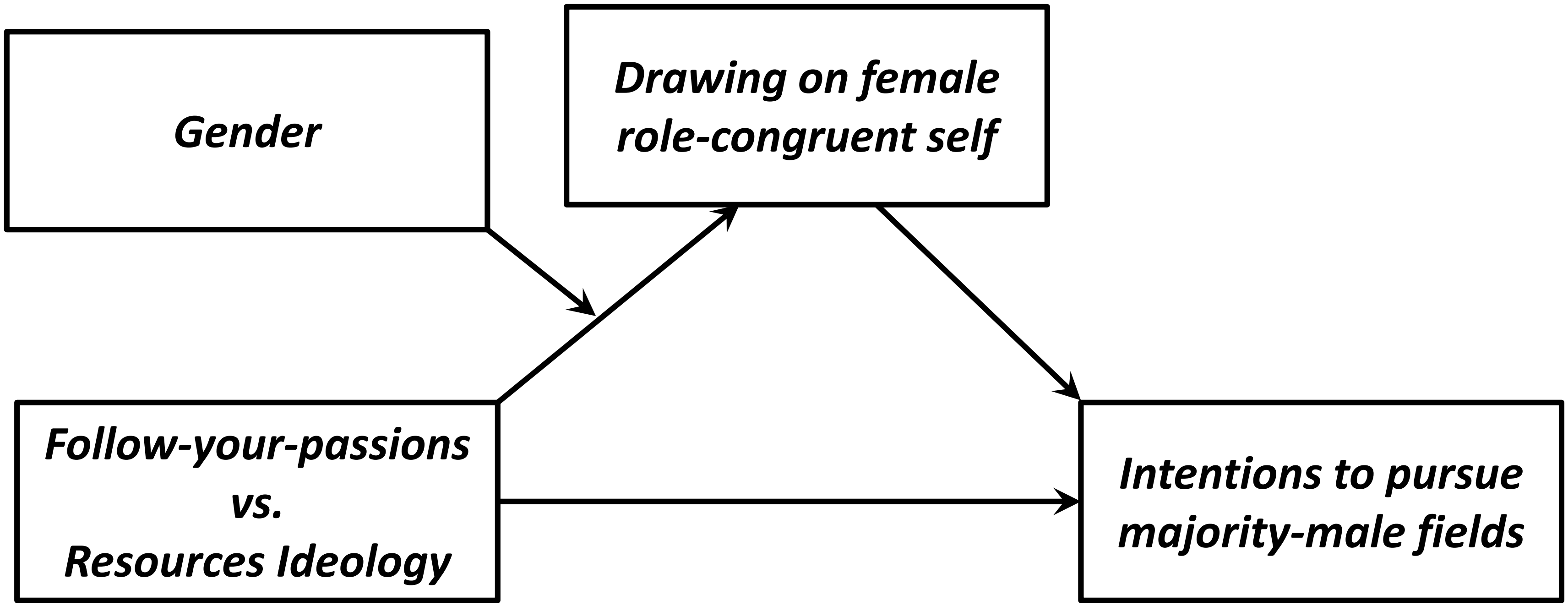
Or

**Does effect of instance on *Y* depend on an individual's *W*?**

Data should be a two-instance repeated-measures design with a **person level covariate**.

# FOLLOW YOUR PASSIONS

**Let's consider a specific part of our theoretical model from the follow-your-passions study:**




We hypothesize the effect of ideology on drawing on the female role-congruent self is moderated by gender (stronger among women).

How can we test this hypothesis, and how can we estimate the effect of ideology for men and for women separately?

Siy, Germano, Vianna, Azpeitia, Yan, Montoya, Cheryan, in press, *JPSP*

# TESTING THE INTERACTION

1. Setup two regression equations, one for each instance

$$Y_{1\mathrm{i}} = b_{10} + b_{11}W_i + \epsilon_{1i}$$

$$Y_{2\mathrm{i}} = b_{20} + b_{21}W_i + \epsilon_{2i}$$

Is $b_{11}$ different from $b_{21}$?

2. Take the difference between those two regression equations

$$Y_{2i} - Y_{1i} = (b_{10} - b_{20}) + (b_{11} - b_{21})W_i + \ (\epsilon_{1i} - \epsilon_{2i}) = b_0 + b_1W_i + \epsilon_i$$

3. Regression weight for person level covariate in Step 2 tests moderation.

Estimate equation above and test if $b_1$ is significantly different from zero

Judd, McClelland, and Smith, 1996

# FOLLOW YOUR PASSIONS

$$Y_{2i} - Y_{1i} = .62 + .48W_i + \epsilon_i$$

The coefficient $\widehat{b_1} = .48$ is statistically significant, meaning that gender significantly moderates the effect of ideology on drawing on your female role-congruent self ($t(670) = 3.96$, $p < .001$).

**In many senses, this raises more questions...**

- What is the effect of ideology for women? Is it significant?
- What is the effect of ideology for men? Is it significant?

Siy, Germano, Vianna, Azpeitia, Yan, Montoya, Cheryan, in press, *JPSP*

# WHAT'S MISSING?

- **Definitions of Conditional Effects**
  - What is the effect of ideology for women, for men?
  - What is the effect of gender in each ideology?
- **Inference about Conditional Effects**
  - **Pick-a-point approach**: Is the effect of ideology significant for women? For men?
  - **Johnson-Neyman**: For what values of my moderator is the effect of ideology significant?
- **How to deal with multiple moderators**
  - **Three-way interactions**: Is the moderation by gender moderated by age?
  - **Multiple two-way interactions**: Is the effect of ideology also moderated by age?

Issues all addressed in Montoya (2019) *Behavioral Research Methods*

Montoya, 2019

# CONDITIONAL EFFECTS IN WITHIN-SUBJECT MODERATION

$$Y_{2i} - Y_{1i} = (b_{20} - b_{10}) + (b_{21} - b_{11})W_i + (\epsilon_{1i} - \epsilon_{2i}) = b_0 + b_1 W_i + \epsilon_i$$

**Given a value of *W* what is the effect of instance on the outcome?**

$Y_{2i} - Y_{1i}$ is a quantification of the effect of instance, which means that the conditional effect of instance $\theta_{X\to Y}(W) = b_0 + b_1 W$

$$\hat{\theta}_{X\to Y}(W) = .62 + .48W_i$$

**Men (Gender = 1)**

$$\theta_{X\to Y}(1) = .62 + .48(1)$$
$$\theta_{X\to Y}(1) = 1.10$$

Men on average draw on their female role-congruent self 1.10 units more in the passions condition than the resources condition

**Women (Gender = 2)**

$$\theta_{X\to Y}(2) = .62 + .48(2)$$
$$\theta_{X\to Y}(2) = 1.58$$

Women on average draw on their female role-congruent self 1.58 units more in the passions condition than the resources condition

Montoya, 2019

# INFERENCE FOR CONDITIONAL EFFECTS

Select a value of the moderator (*W*)

$$\theta_{X \to Y}(W) = b_0 + b_1 W$$

The standard error of $\theta_{X \to Y}(W)$ is

$$s_{\theta_{X \to Y}(W)} = \sqrt{(s_{b_0}^2 + 2W s_{b_0 b_1} + W^2 s_{b_1}^2)}$$

Squared standard error of $b_0$

Covariance of $b_0$ and $b_1$

Squared standard error of $b_1$

The ratio of the effect to standard error is *t*-distributed as $t(df_{residual})$ under the null hypothesis that the effect of instance is zero at that moderator value.

$$\frac{\hat{\theta}_{X \to Y}(W)}{\hat{s}_{\hat{\theta}_{X \to Y}(W)}} \sim t_{df}$$

Montoya, 2019

# FOLLOW YOUR PASSIONS

This model can be estimated in MEMORE

```
memore y = gendSpas gendSinc /w = gendr /model = 2 /plot = 1.
```

***Conditional effect of ideology:***

Men on average draw on their female role-congruent self 1.10 units more in the passions condition than the resources condition, **and this effect is statistically significant (*t*(670) = 12.15, *p* < .01).**

Women on average draw on their female role-congruent self 1.58 units more in the passions condition than the resources condition, **and this effect is statistically significant (*t*(670) = 19.61, *p* < .01).**

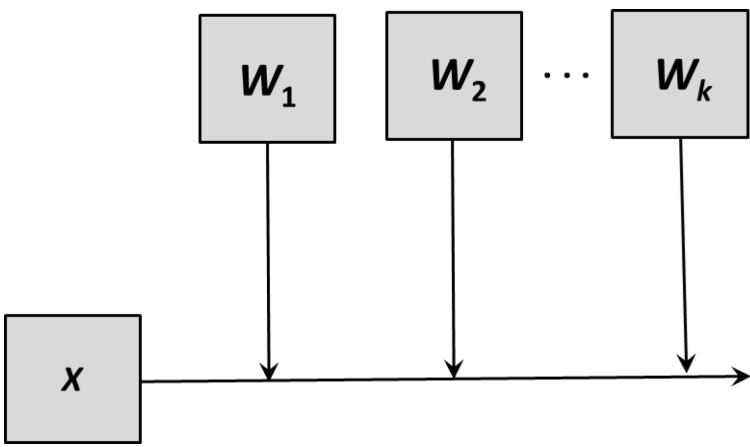

Siy, Germano, Vianna, Azpeitia, Yan, Montoya, Cheryan, in press, *JPSP*

# MULTIPLE MODERATOR MODELS

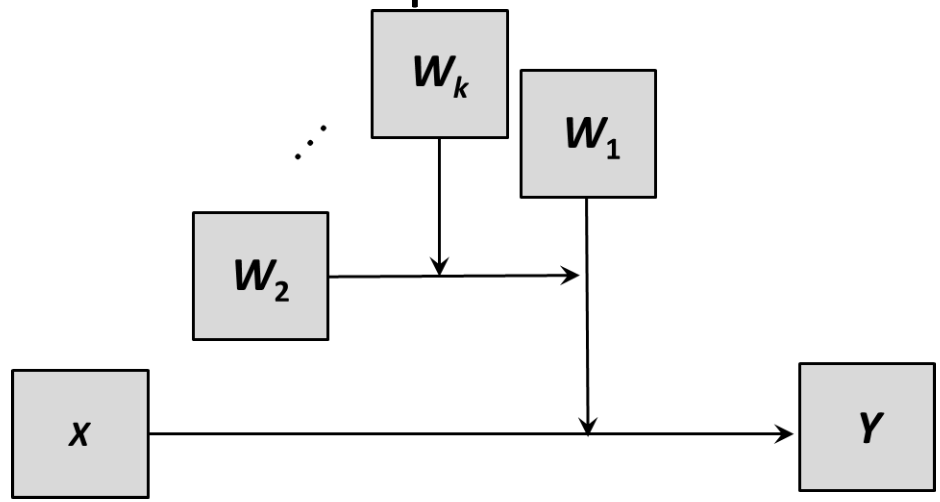
Multiple moderator models are also included!

- Can have up to 5 moderators

**Model 2: Additive Moderation**

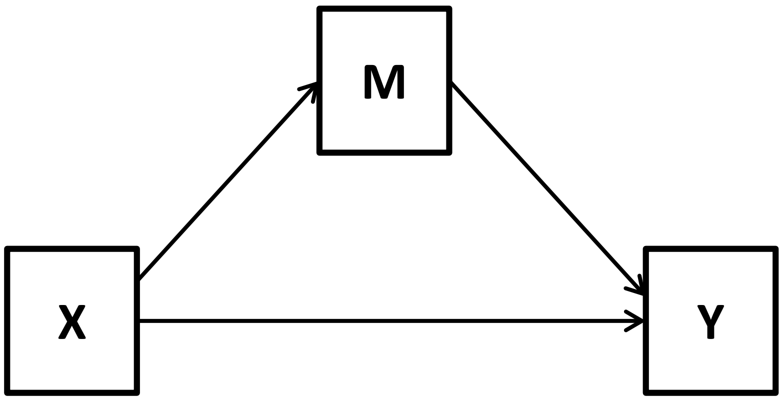


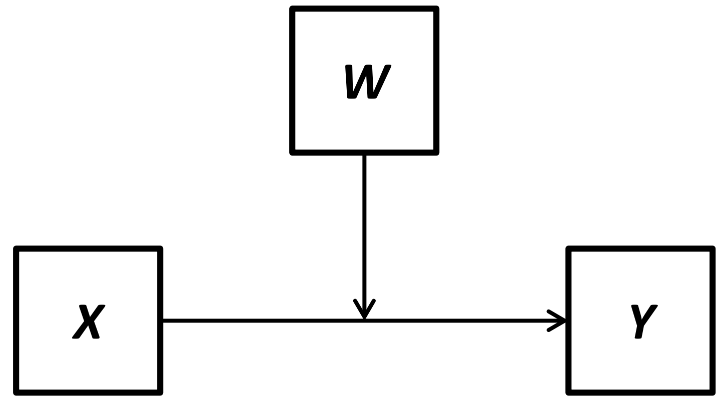

**Model 3: Multiplicative Moderation**




Montoya, 2019

# COMBINING MEDIATION AND MODERATION: CONDITIONAL PROCESS ANALYSIS




Research questions:

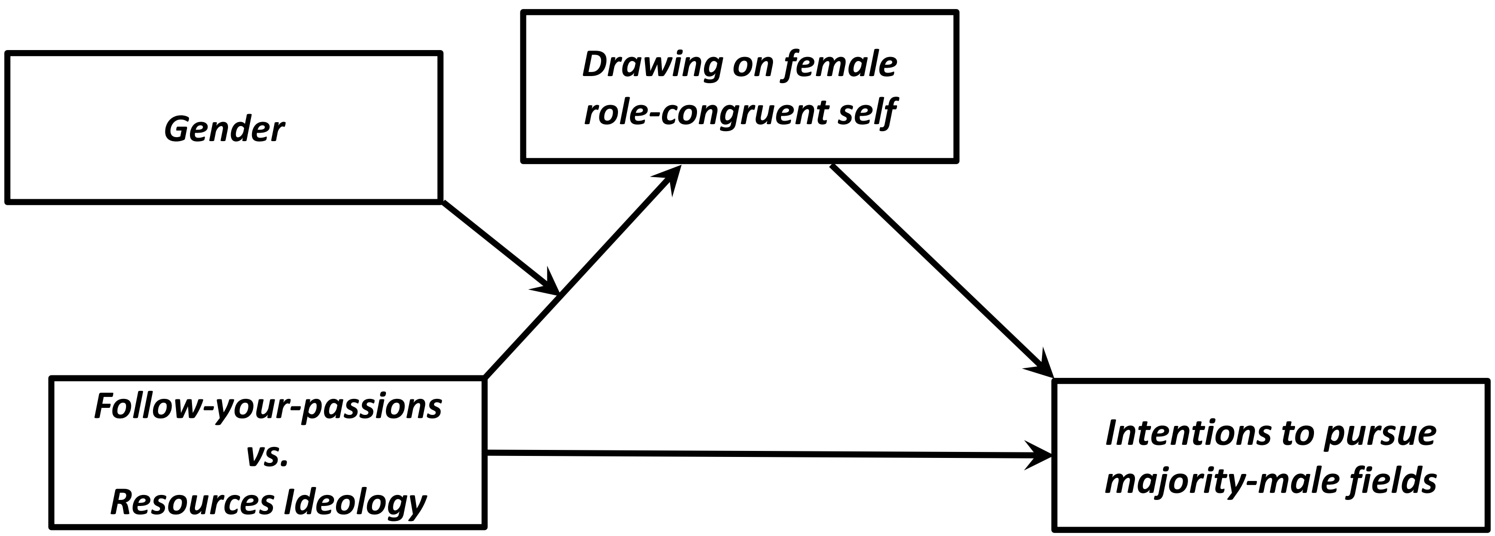
- Does the process through which *X* affects *Y* through *M* depend on *W*?
- Are there certain groups where *X* affects *Y* through *M* and certain groups where this process does not occur?

**Conditional process analysis** allows a mediated process to be moderated. Now the indirect effect can be defined as a *function of the moderator.*

# CPA IN TWO-INSTANCE REPEATED-MEASURES DESIGNS

Extending the path analytic from Montoya & Hayes (2017) we can now allow for moderation of a mediated pathway.

**First stage moderated mediation** allows *W* to moderate the path between the within-subjects factor and the mediator.




Montoya, *in prep (https://psyarxiv.com/wrtb6)*

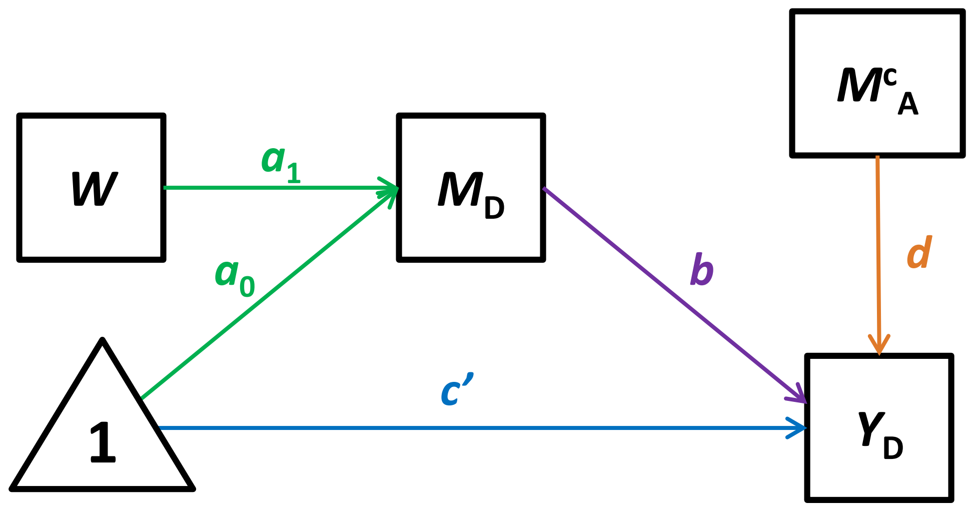
# EQUATIONS AND PATH DIAGRAM

**First stage moderated mediation** allows *W* to moderate the path between the within-subjects factor and the mediator.

$$M_{2i} - M_{1i} = a_0 + a_1 W_i + \epsilon_{Mi}$$

$$\theta_{X \to M}(W) = a_0 + a_1 W_i$$

$$Y_{Di} = c' + bM_{Di} + dM_{Ai}^{c} + \epsilon_{Yi}$$




What is the indirect effect?

$$\theta_{X \to M}(W) \times b = (a_0 + a_1 W)b = a_0 b + a_1 bW$$

Indirect effect is a *function* of the moderator

Montoya, *in prep (https://psyarxiv.com/wrtb6)*

# MAKING INFERENCE

$$\theta_{X\rightarrow M}(W) \times b = (a_0 + a_1 W)b = a_0 b + a_1 bW$$

**Conditional Indirect Effects**
Select a value of *W*, plug that into the equation for the indirect effect, and use bootstrapping to make inference about the indirect effect at that value

Does the indirect effect *depend* on the moderator?

If $a_1 b = 0$ then the indirect effect *does not* depend on W

$$\theta_{X\rightarrow M}(W) \times b = a_0 b + 0 * W = a_0 b$$

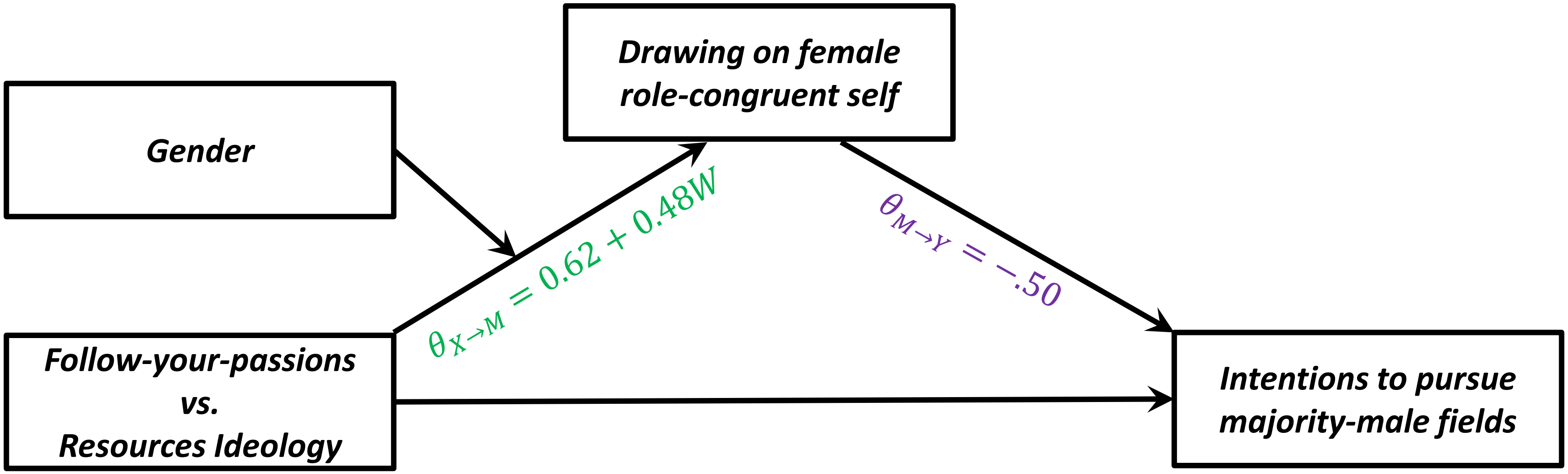
$a_1 b$ can be called the **index of moderated mediation**

A test on the index will indicate if the indirect effect depends on *W.* We can do this formal test using bootstrapping.

Montoya, *in prep (https://psyarxiv.com/wrtb6)*

# FOLLOW YOUR PASSIONS

```
memore y = FMpas FMinc /m = gendSpas gendSinc /w =
gendr /model = 14.
```



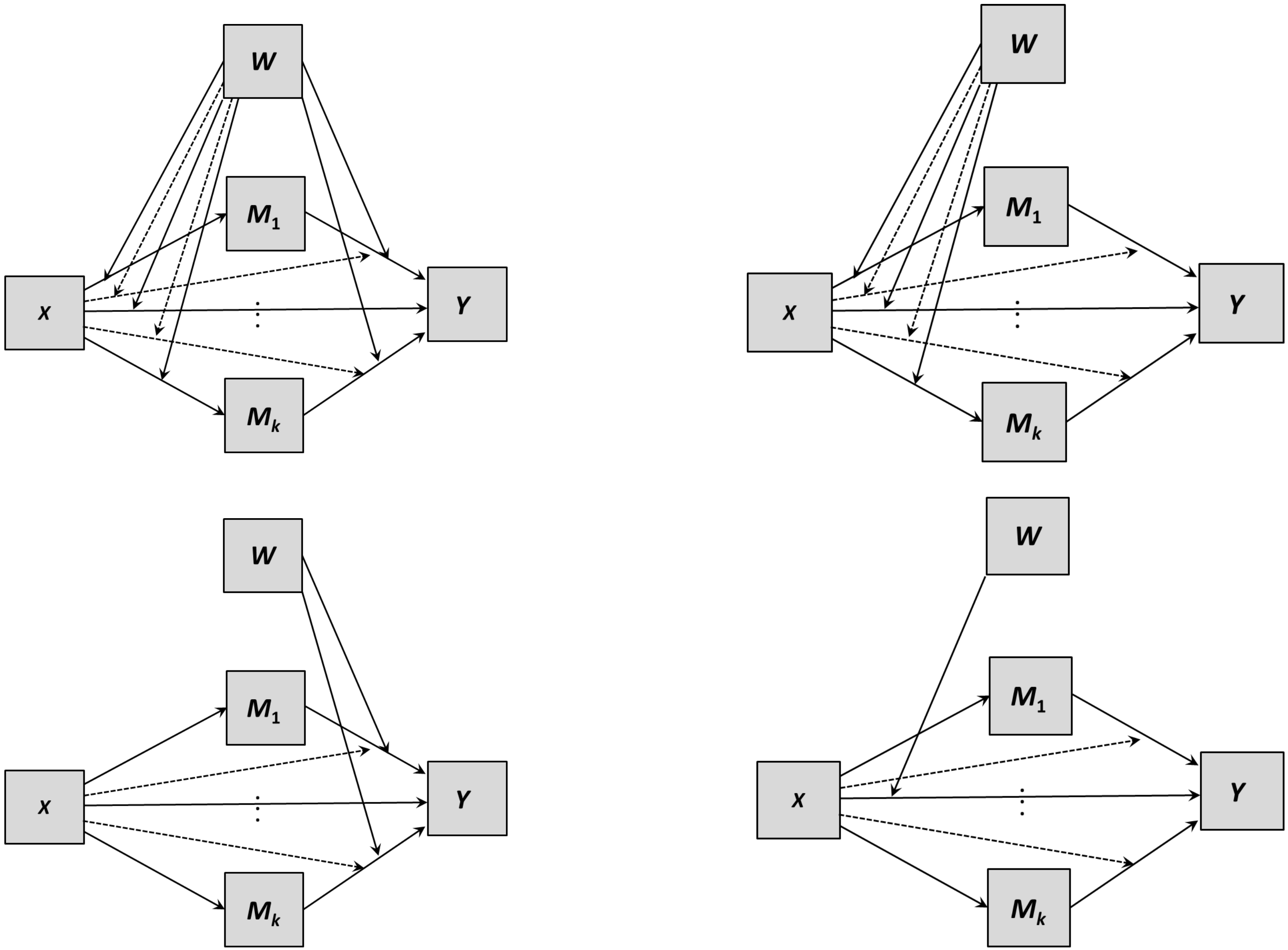

The indirect effect for both men and women was such that **the follow-your-passions ideology decreased interest through drawing on feminine self** (Men: -.55 [-.66, -.44], Women: -.78 [-.94, -.63]).

The *index of moderated mediation* was significantly different from zero (-.24 [-.37, -.12]), meaning the **indirect effect through drawing on the feminine self was stronger for women than for men**.

Siy, Germano, Vianna, Azpeitia, Yan, Montoya, Cheryan, in press, *JPSP*

# MEMORE V3: MODELS 4 - 18

The latest version of MEMORE has expanded to models with a single moderator on any combination of paths in the mediation.




Montoya, *in prep (https://psyarxiv.com/wrtb6)*

# FUTURE DIRECTIONS

- Formal release of MEMORE with final publication
- MEMORE for R
- Generalizations similar to those described for mediation
  - More than two conditions
  - Dichotomous outcomes (and other GzLMs)
- Causal inference in moderated mediation (between and within)
- Building bridges to more complex models:
  - Multilevel models
  - Latent growth curve models
  - Latent Difference Score
  - Diff-in-Diff

**These areas have had very limited discussion of conditional process analysis**

# OVERVIEW

**Two-Instance Repeated-Measures Designs**

- Mediation Analysis
- Moderation Analysis
- Conditional Process Analysis

**Meta-Science**

- Quality of Evidence in Mediation Analysis
- Sample Size Planning in Mediation and Moderated Mediation
- Registered Reports

**A Call to Arms**

# ANALYTICAL FLEXIBILITY IN MEDIATION AND MODERATED MEDIATION

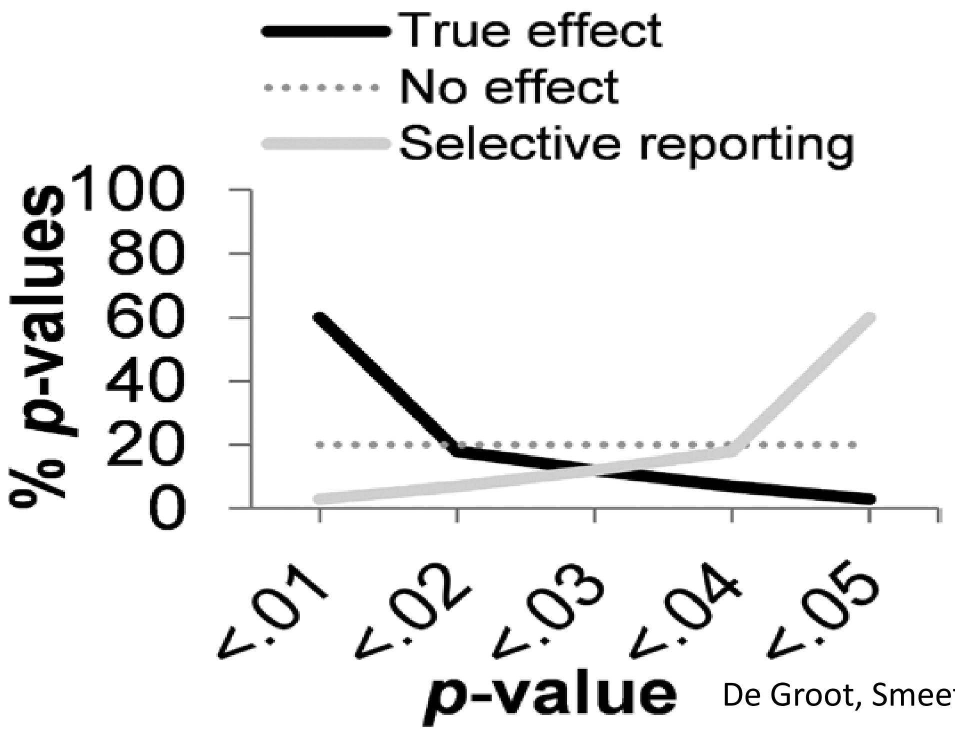
- Easy to swap around variables in different orientations
  - Mediator/Outcome
  - Covariate/Moderator
- Easy to move moderation to different paths
  - First-stage vs. Second-stage, Moderation of Direct Effect
  - Additive or Multiplicative
- Multiple options for inferential methods
- Multiple options of estimation methods
- Exploitation of randomness in resampling procedures



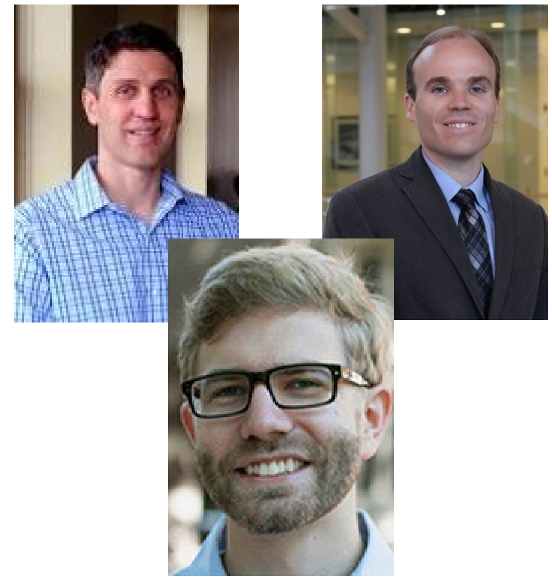
Charlton, Montoya, Price, Hilgard (*under review*), https://psyarxiv.com/ck2r5/

# *P*-CURVE FOR MEDIATION

*P*-curve is a method of analyzing *p*-values from a set of literature to investigate if there is evidence of *p*-hacking




De Groot, Smeets, 2017



Mediation analyses are most commonly reported with a bootstrap CI

**How can we evaluate the presence of *p*-hacking for mediation analyses?**

Charlton, Montoya, Price, Hilgard (*under review*), https://psyarxiv.com/ck2r5/

# *P*-CURVE FOR MEDIATION

Relative Proximity measures the distance of the CI from zero, relative to it's width

If CI does not contain zero:

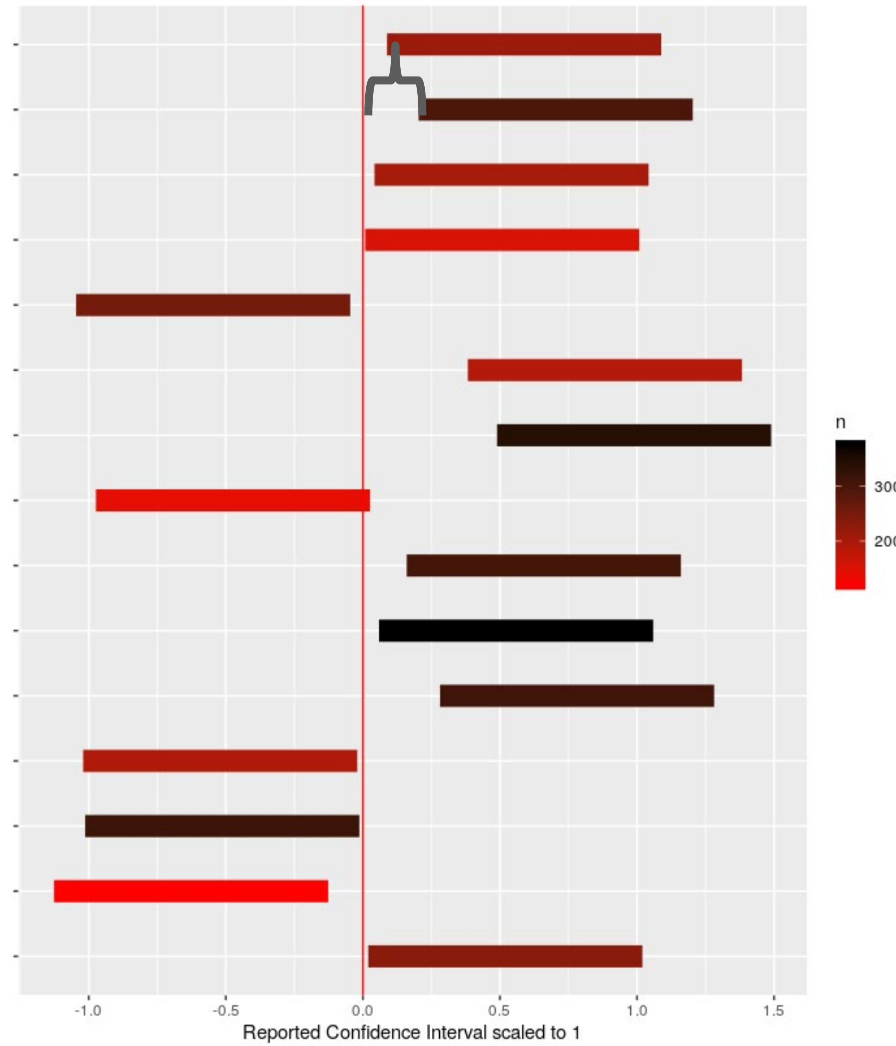
$$\frac{\min(|CI_{2.5\%}|,|CI_{97.5\%}|)}{|CI_{2.5\%} - CI_{97.5\%}|}$$

If CI contains zero:

$$-1\frac{\min(|CI_{2.5\%}|,|CI_{97.5\%}|)}{|CI_{2.5\%} - CI_{97.5\%}|}$$

If RP is negative: no evidence
If RP is close to zero: limited evidence
If RP is far from zero: strong evidence

Relative Proximity




Gotz, O'Boyle, Gonzales-Mule, Banks (2021) *Psychological Bulletin*
Charlton, Montoya, Price, Hilgard (*under review*), https://psyarxiv.com/ck2r5/

# *P*-CURVE FOR MEDIATION

Gotz, O'Boyle, Gonzales-Mule, Banks, 2021, *Psychological Bulletin*

Using a sample from APA journals, confidence intervals for mediation analyses

- Small proportion of null results suggests publication bias
- Large piling at zero suggests "CI"-hacking or low power & file drawer

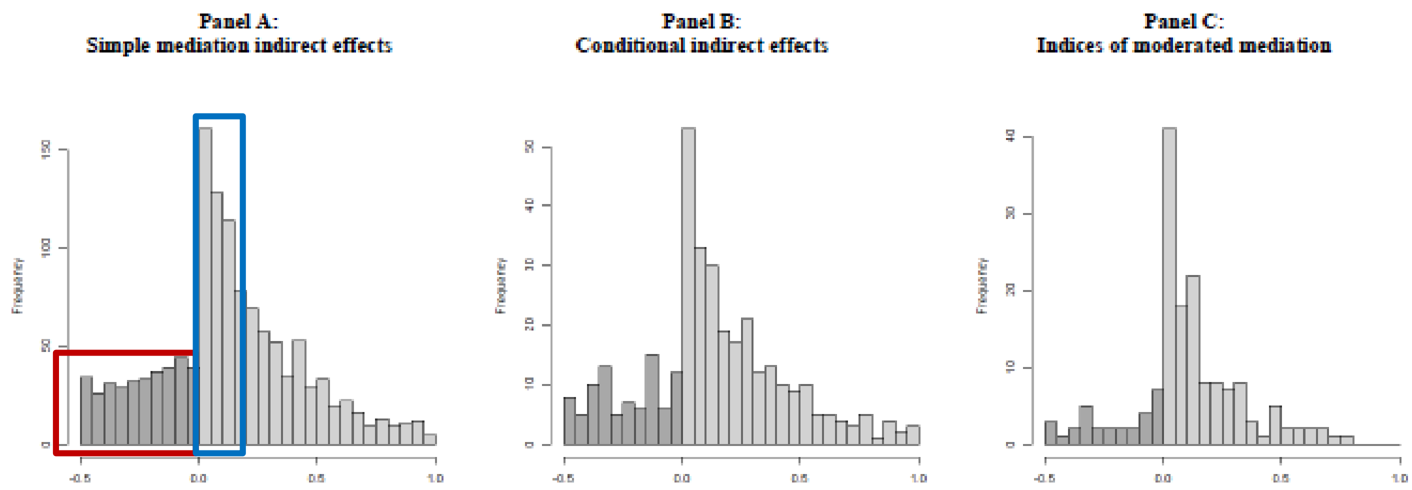



*Figure 5.* Relative proximities of hypothesized tests of mediation with indirect effects (Panel A, k = 1,372), conditional indirect effects (Panel B, k = 361), and indices of moderated mediation (Panel C, k = 161). Bars in dark grey represent frequencies of non-significant point estimates, whereas bars in light grey frequencies of significant ones.

Charlton, Montoya, Price, Hilgard (*under review*), https://psyarxiv.com/ck2r5/

# MARKETING PAPERS (2017 – 2020)

Keywords: mediat*, indirect, Hayes, and *PROCESS*


Journal of Consumer Research

79 papers


Journal of Consumer Psychology

120 papers


Journal of Marketing Research

66 papers


Journal of Marketing

43 papers

# SUMMARY OF RESULTS

**N = 308**

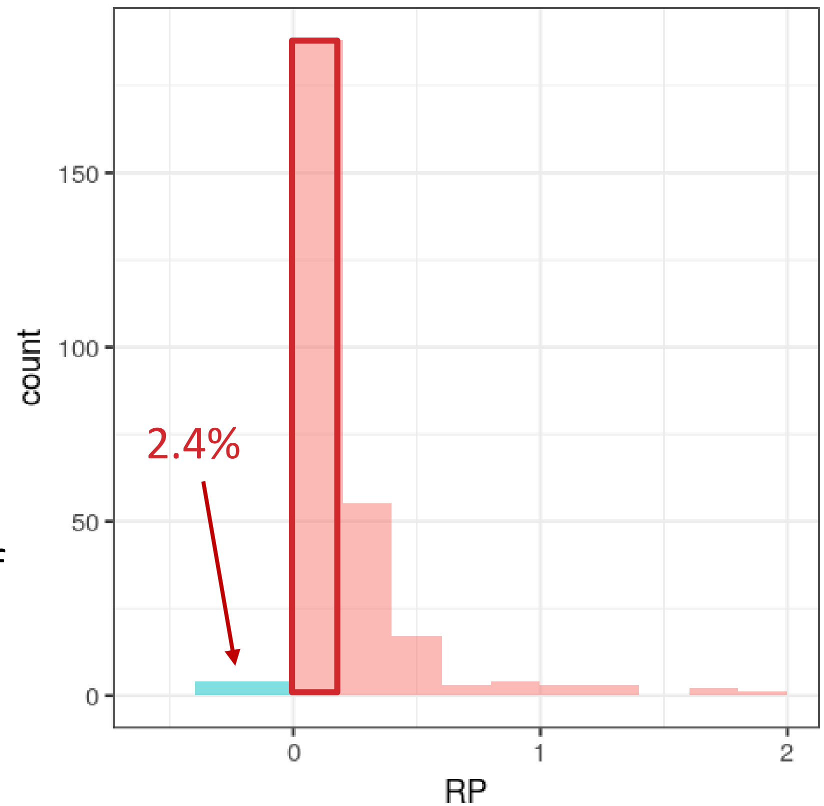
**Type of Estimate**

- 212 indirect effect (69%)
- 22 conditional indirect (7%)
- 73 index of moderated mediation (24%)

**Sample Size**

- Min: 56, Max: 61,642
- Mean: 712, Median: 212.5
- Missing: 1

There is an overwhelming stack of results very close to zero




This suggests potential issues with the quality of evidence in mediation analysis.

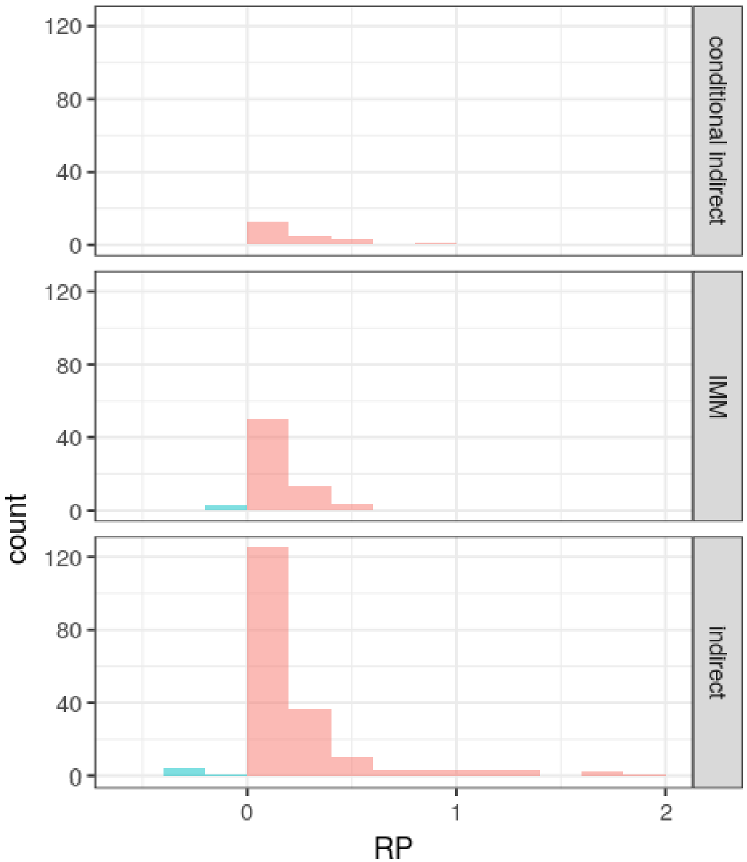
"CI-hacking" or Low Power and File Drawer

Charlton, Montoya, Price, Hilgard (*under review*), https://psyarxiv.com/ck2r5/

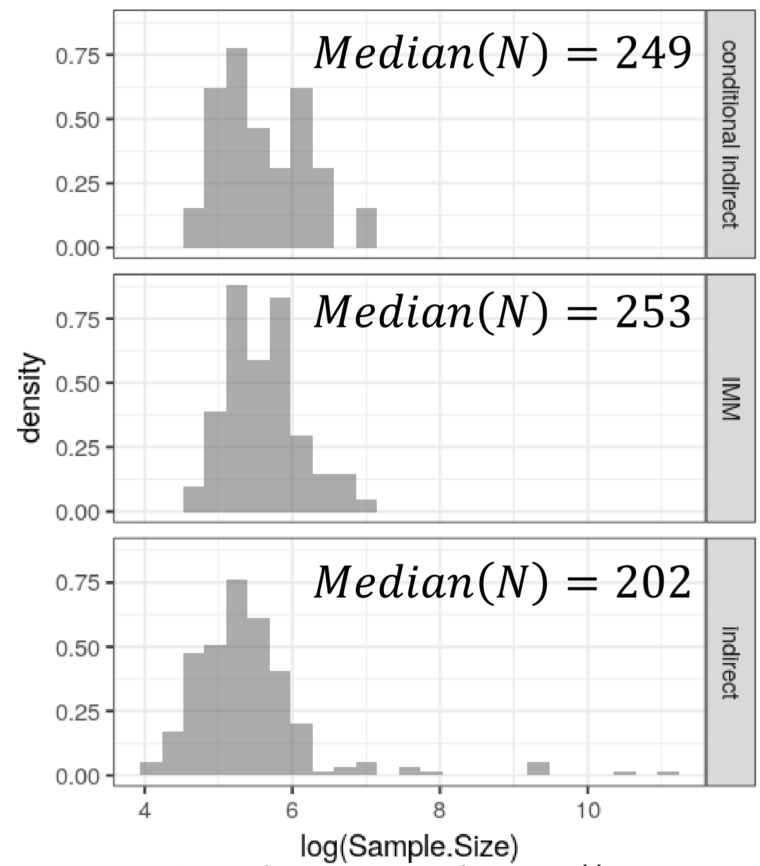
# TYPES OF ESTIMATE AND SAMPLE SIZE

Larger samples **should** be needed to detect indices of moderated mediation compared to indirect effects, so the relative proximity distribution should look different or the sample sizes should be different.

**RP does not significantly differ across Type of Estimate.** $F(2, 281) = 1.24$, $p = 0.29$




**Log(N) does not significantly differ across Type of Estimate** $F(2, 291) = 1.14$, $p = 0.32$

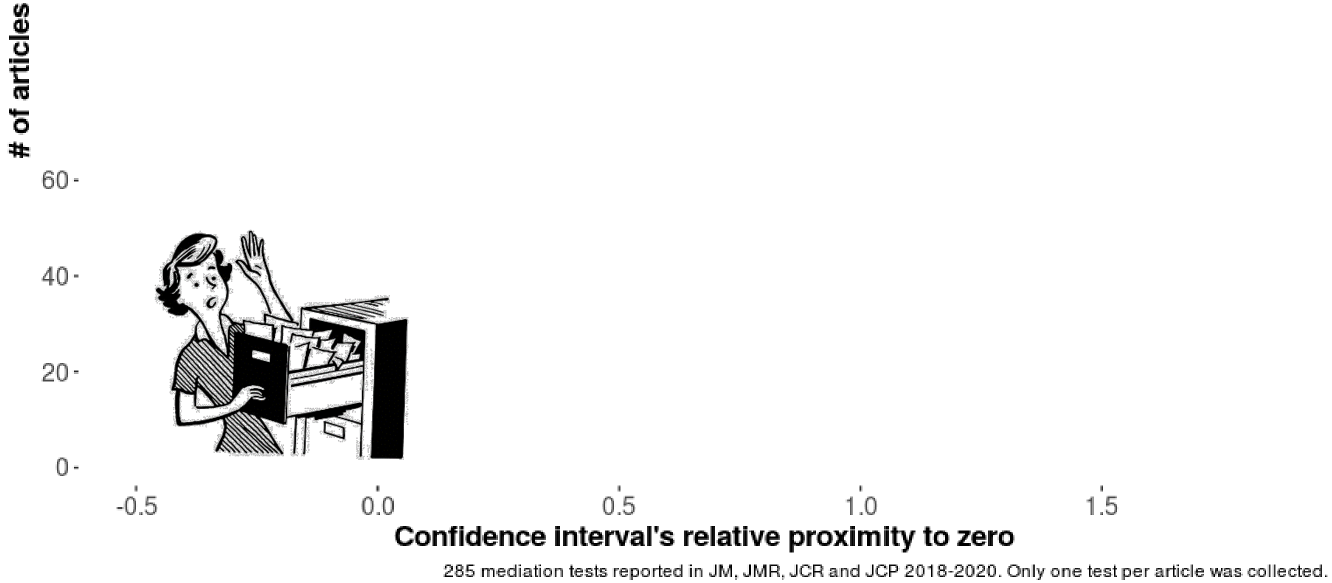



Charlton, Montoya, Price, Hilgard (*under review*), https://psyarxiv.com/ck2r5/

# LOW POWER?

**We conducted simulations to evaluate what power curves fit the data best.**

If the results are due to low power, this indicates another problem: A very large file drawer.




Charlton, Montoya, Price, Hilgard (*under review*), https://psyarxiv.com/ck2r5/

# P-CURVE WITH MEDIATION: SUMMARY

Evidence from this study and Gotz et al. (2021) in psychology suggest there are potential issues with bias in reporting and/or conduct on mediation analysis

- “CI-hacking” to get confidence intervals just over zero
- Low power & file drawer

Researchers do not seem to be adjusting their sample size to account for complexity of their model.

Solutions?

- **Improve power**: Stronger manipulations, Larger samples, Better measures
- **Preregistration:** Make a record of your planned analyses before data collection
- **Replication:** Most replication attempts have focused on primary hypotheses (main effects), but we also need to replicate the mediation

Charlton, Montoya, Price, Hilgard (*under review*), https://psyarxiv.com/ck2r5/

# SAMPLE SIZES FOR MODERATED MEDIATION

Previous research in mediation has suggested *N* = 500 to detect “medium” indirect effects, and now we are looking at moderation of indirect effects (Fritz & MacKinnon, 2007).

Team of 3 Undergraduates + Jessica coded more than 400 published examples of moderated mediation models

**Sample Size**

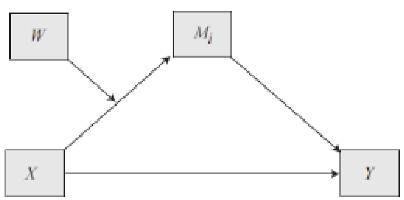
Median sample size: 285

# WHAT MODELS ARE PEOPLE USING?

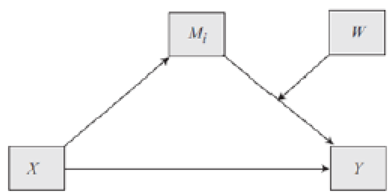
Team of 3 Undergraduates + Jessica coded more than 400 published examples of moderated mediation models

**Model Specification**

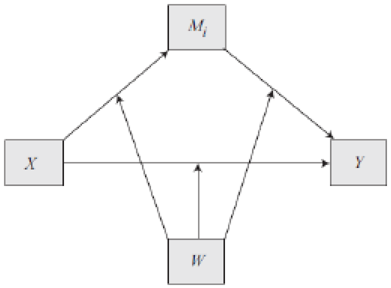
*N* = 128



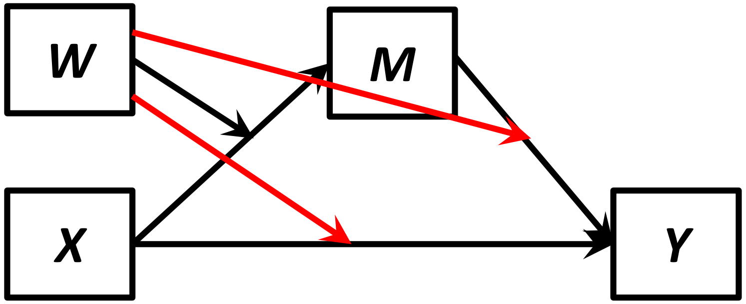

*N* = 74




*N* = 57




**Follow up:** How does specification impact power?



# A TOOL FOR FINDING EXAMPLES

A common question our lab receives is if we know of an example of a paper that uses Model XX from PROCESS.

Jessica created a database from her literature search, and has continued to add to it through the help of undergrad RAs. You can also submit a form to add your paper to the database if you want others to be able to find it!

https://www.jlfossum.com/home/moderated-mediation-article-database

Model Number | Type of Manipulated Variable | Number of Moderators | Number of Mediators | Year

Estimation Procedure | Research Area | Journal | Number of X Variables | Number of Y Variables

| Article | Journal | First Author | Research Area | Model Number | Sample Size | Estimation Procedure | Number of X Variables | Type of X Variable | Number of Y Variables | Mediators | Moderators | Covariates |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| "They've conspired against us": Understanding the role of social identification and conspiracy beliefs in justification of ingroup collective behavior | EUROPEAN JOURNAL OF SOCIAL PSYCHOLOGY | Chayinska, M | Psychology | 7 | 315 | Regression | 1 | | 1 | 1 | 1 | 0 |
| A Conditional Process Analysis of the Teacher Confirmation-Student Learning Relationship | COMMUNICATION QUARTERLY | Goldman, ZW | Communication | 7 | 208 | Regression | 1 | | 1 | 1 | 1 | 1 |
| A Latent Growth Moderated Mediation Model of Math Achievement and Postsecondary Attainment: Focusing on Context-Invariant Predictors | JOURNAL OF EDUCATIONAL PSYCHOLOGY | Guglielmi, RS | Psychology | | 8791 | Regression | 1 | | 1 | 3 | 2 | 2 |
| A Longitudinal Study of Inhibited Temperament, Effortful Control, Gender, and Anxiety in Early Childhood | CHILD & YOUTH CARE FORUM | Niditch, LA | Psychology | 59 | 1226 | SEM | 1 | | 1 | 1 | 1 | 4 |
| A Moderated Mediation Model for Board Diversity and Corporate Performance in ASEAN Countries | SUSTAINABILITY | E-Vahdati, S | Science & Technology - Other Topics | 58 | 264 | SEM | 2 | | 1 | 1 | 1 | 3 |
| A Social Influence Interpretation of Workplace Ostracism and Counterproductive Work Behavior | JOURNAL OF BUSINESS ETHICS | Yang, J | Business & Economics | 21 | 156 | Regression | 1 | | 1 | 1 | 2 | 4 |
| A genetic variant brain-derived neurotrophic factor (BDNF) polymorphism interacts with hostile parenting to predict error-related brain activity and thereby risk for internalizing disorders in children | DEVELOPMENT AND PSYCHOPATHOLOGY | Meyer, A | Psychology | 7 | 201 | Regression | 1 | | 1 | 1 | 1 | 2 |
| A pre-event assessment of | TOURISM | Tournois, L | Environmental | 9 | 317 | Regression | 1 | | 1 | 1 | | |

# PREREGISTRATION OF MEDIATION ANALYSIS

**An example of a "preregistered" mediation analysis:**

"We plan to conduct mediation tests (including multiple mediation) using the MEMORE macro (Montoya & Hayes, 2017) and the PROCESS macro (Hayes, 2017)."

**We need better guidelines for what decisions to include in our preregistrations:**

- Planned sample size
- $\alpha$-level/CI-level
- Role of different variables in the analysis (e.g., condition, mediator, outcome, covariate), and how they are computed
- Estimation Method
- Inferential method and any important specifications for that method (e.g., how many bootstraps, seed number)
- Plans for sensitivity analysis: Benchmarks for acceptable correlations for evaluating confounding

Montoya (in press) *Handbook of Research Methods in Social and Personality Psychology*

# IS PREREGISTRATION ENOUGH?

**Limitations of preregistration**

- A plan is not always a *good* plan
- Review process can conflict with prereg
- Does not address publication bias/file drawer

**Registered Reports: An alternative**

- Submit an introduction, methods, and analysis plan to a journal (Stage 1)
- Stage 1 submission is peer-reviewed (can be revised)
- In-principle acceptance: journal will publish regardless of the results
- Researchers collect data, write up results and discussion (Stage 2)
- Stage 2 submission is peer-reviewed for adherence to original plan

# IMPLEMENTATION

Registered reports are being increasingly adopted across scientific disciplines.


More than 350 journals accept registered reports!

Peer Community In Registered Reports can also allow you to do an RR outside a journal

# A NEW DATABASE FOR REGISTERED REPORTS

Being implemented as go-to tool with Open Science Framework

Easy-to-search database with all journals adopting registered reports

Provides information about policies, academic field, etc.

Scientific Discipline | 5 Year Range | Registered Replications | Secondary Data Analysis allo... | H-Index Range

Special Issue | Novel studies within RR | Multiple Studies | Word Limits | Research Area

| Journal Hyperlink | Scientific Discipline | Research Area | h-index | Power Requirements | Word Limits | Novel studies within RR | 5 year IF | 2017 IF (Two Year) |
|---|---|---|---|---|---|---|---|---|
| AAS Open Rese... | General | null | null | 0.9 or high... | No mention | No mention | None | null |
| AERA Open | Psychology | Learning; Devel... | null | No mention | 8,000 words | Yes | None | null |
| AIMS Neuroscie... | Psychology | Biopsychology/... | 6 | 0.9 or high... | None | Yes | None | 1.028 |
| AMRC Open Re... | General | null | null | 0.9 or high... | No mention | No mention | None | null |
| Academia Jour... | Medicine | null | null | No mention | No mention | No mention | None | null |
| Acta Psychologi... | Psychology | null | 88 | 0.9 or high... | No mention | Yes | 2.258 | 1.632 |
| Adaptive Huma... | Psychology | null | 13 | No mention | 1,500 extended ... | Yes | None | 3.326 |
| Addiction Rese... | Psychology | null | 45 | 0.9 or high... | No mention | Yes | 1.94 | 2.478 |
| Advances in Me... | Psychology | Research Meth... | null | No | Soft word limits... | Yes | N/A | null |
| Affective Science | Psychology | null | null | No mention | 2,000 words (Int... | No mention | None | null |
| Alimentary Pha... | Medicine (gastr... | null | 155 | No mention | 12-20 pages dou... | No mention | None | 7.357 |
| American Journ... | Political Science | null | 136 | No mention | 10,000 words | Yes | None | 5.22 |
| American Politi... | Political Science | null | 148 | No mention | 12,000 words | Yes | 6.224 | 3.252 |
| American Politi... | Political Science | null | 41 | No mention | No mention | Yes | 1.444 | 1.089 |

| Journal Hyperlink | Multiple Studies | Access to Data Required | Access to Material Required | Changes to Introduction Allowed | Deadline after Stage 1 approval | Special Issue | Exploratory Analyses allowed(not in original plan) | External Preregister required? | Preliminary Studies allowed? |
|---|---|---|---|---|---|---|---|---|---|
| AAS Open Rese... | Incremental Re... | Yes | No mention | Only minor styli... | Request specifi... | No | Yes, seperate section | No mention | Yes |
| AERA Open | No mention | No, but encoura... | No, but encoura... | No mention | Date specific | Yes | Yes | No, but enco... | No menti... |
| AIMS Neuroscie... | Yes, sequential ... | Yes | Yes | No | Request specifi... | No | Yes, separate section | No mention | Yes |
| AMRC Open Re... | Incremental Re... | Yes | No mention | Only minor styli... | Request specifi... | No | Yes, seperate section | No mention | Yes |
| Academia Jour... | No mention | No mention | No mention | No mention | No mention | No | No mention | No mention | No menti... |
| Acta Psychologi... | Incremental Re... | Yes | Yes | Only minor styli... | Request specifi... | No | Yes, seperate section | No mention | Yes |
| Adaptive Huma... | No mention | Encouraged | No mention | No mention | Out of date | Yes | No mention | No mention | No menti... |
| Addiction Rese... | Incremental Re... | Yes | Yes | Only minor styli... | Request specifi... | No | Yes, seperate section | Yes | Yes |
| Advances in Me... | Incremental Re... | Yes | Yes | No | No mention | No | Yes, separate section | Yes | Yes, not m... |
| Affective Science | No mention | Yes | Yes | No mention | No mention | No | No mention | No mention | No menti... |
| Alimentary Pha... | No mention | No, but encoura... | No, but encoura... | No mention | No mention | No m... | No mention | No mention | No menti... |
| American Journ... | No mention | Yes, data provid... | Yes | No mention | Date specific | Yes | Yes, separate section | Yes | No menti... |
| American Politi... | No mention | Yes, data provid... | Yes | No mention | Date specific | Yes | Yes, separate section | Yes | No menti... |
| American Politi... | No mention | Yes, data provid... | Yes | No mention | Date specific | Yes | Yes, separate section | Yes | No menti... |

1 - 250 / 250

| Journal | OSF Link 1 | OSF Link 2 | OSF Link 3 |
|---|---|---|---|
| AAS Open Research | Author Guidelines | null | null |
| AERA Open | Call for submissions | null | null |
| AIMS Neuroscience | Editorial | Guidelines | null |
| AMRC Open Research | Author Guidelines | null | null |
| Academia Journal of Stroke | null | null | null |
| Acta Psychologica | Guidelines | null | null |

1 - 243 / 243

Data Last Updated: 8/21/2020 11:00:53 AM | Privacy Policy

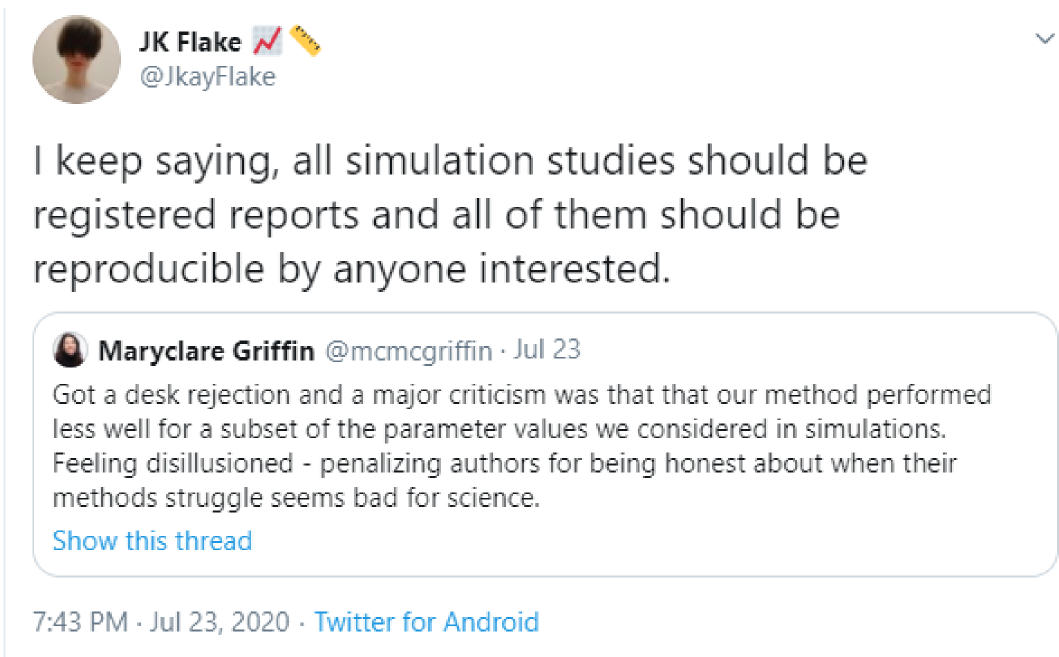
Montoya, Krenzer, Fossum, 2021

# HOW REGISTERED REPORTS ALIGN WITH THE VALUES OF QUANTITATIVE PSYCHOLOGISTS?

Simulation studies could easily be transitioned to a registered report model

Could reduce wasted time/resources from rerunning sims

We could also benefit from providing open code/data



# DOES SIMULATION RESEARCH HAVE A REPLICABILITY PROBLEM? REPLISIMS

**Goal: Generate replications of "high impact" simulation studies that are relatively recent, which are…**

- Reproducible
- Extendable
- Implement best practices in simulation research

**Leadership Team: Anna Lohmann, Rolf Groenwold, & Kim Luijken**

- QRClab contributing replication studies
- You can also contribute, check out replisims.org

**My Hypothesis: Reproducibility (same data/code different analyst) should be close to perfect, replicability (read the methods and try to repeat sim) could be better**

  - We likely have our own "researcher degrees of freedom" we need to consider

# META-SCIENCE SUMMARY

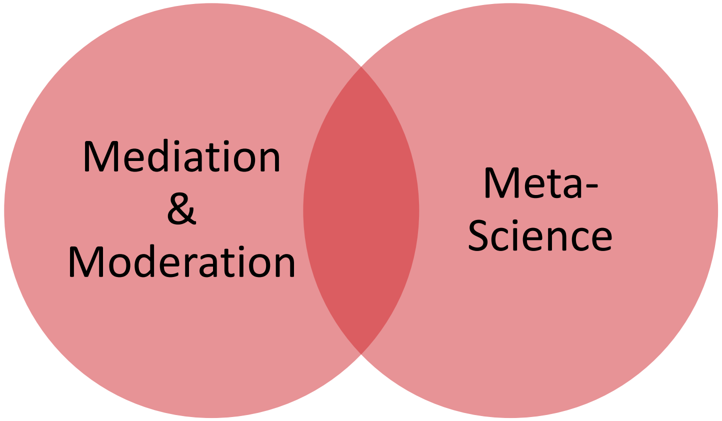
Research in mediation and moderated mediation may have low evidential value.

- Potential evidence of “CI-hacking”
- Low power and a large file drawer
- Undisclosed flexibility in analysis choices

Our team focuses on helping research plan their work better:

- Preregistration (Montoya, in press)
- Power analysis (Montoya, 2022; Fossum & Montoya, under review)
- Registered reports (Montoya, Krenzer, Fossum, 2021)

# SUMMARY & A CALL TO ARMS




**Mediation and moderation** are very important to answering research questions in psychology

- Repeated-measures designs are common but underexplored in this area
- Extension to complex outcomes are underway

Methods development needs to focus on **replicability related practices**

- Current research suggests that sample sizes for mediation and moderated mediation are too low
- Guidelines for preregistration are underdeveloped
- Registered reports offer benefits above and beyond preregistration

# CAN META-SCIENCE HELP QUANTITATIVE PSYCHOLOGY?

**Large open datasets**

- Estimates from replications serve as samples from a sampling distribution

**Encouraging new methods of data visualization**

**Use meta-science approaches to inform quantitative research**

- Consider survey research to understand common “misconceptions”
- Improve understanding of how methods are used “in the wild”?
- Inform simulation parameters

# HOW CAN QUANTITATIVE PSYCHOLOGY CONTRIBUTE TO THE CREDIBILITY REVOLUTION?

**Many open questions in meta-science!**

- What constitutes replication? (*p*-values, effect-size)
- What is the best way to select sample size to improve replicability without wasting too many resources?
- Combining information across studies (IDA, Meta-analysis)
- Methods for detecting data falsification and *p*-hacking

**Preregistration and reporting standards**

- Current guidelines are not well-established for complex analytical techniques

**Improved measurement practices**

**Data visualization**

# LOOKING TO GET INVOLVED?

**Join/attend Society for the Improvement of Psychological Science**

**Contribute to RepliSims (replisims.org)**

**Common journals:**

- *Advances in Methods and Practices in Psychological Science*
- *Collabra* (Official SIPS journal)
- *Meta-Psychology*

**Post pre- and post-prints to PsyArXiv.com and EdArXiv.org**

**Check out the Center for Open Science's Open Science Framework (osf.io)**

- Preregistration
- Open Data / Materials
- Project management

**Join the Twittersphere**

# THANK YOU!

I am available for questions now and in the future via email at akmontoya@ucla.edu

MEMORE can be downloaded from akmontoya.com

RR database can be found at bit.ly/RRdatabase

Slides available at
github.com/akmontoya/USC2023

# PROPENSITY SCORE RECRUITMENT

Dr. Sumner's R01 involves recruiting a group of individuals with PTSD and a Trauma Exposed (non-PTSD) control group

**Goal:** Develop a system for determining whether to recruit individuals on a rolling basis, where we end up with groups that are approximately matched on 8 demographic characteristics (continuous and categorical)

**Methods:** Rolling propensity score matching. Participants are divided into 5 groups based on estimated propensity score. Up to 3 participants are "held out" in each group, allowing for more potential for matching with limited participant waste. We are also conducting simulation studies to evaluate the effectiveness of this process.

**Results:** We have interviewed 69 participants, enrolled 51

Cleveland, Reed, Thomas, Ajijola, Ebrahimi, Hsiai, Lazarov, **Montoya**, Neria, Shimbo, Wolitzky-Taylor, Sumner (2021)

# NUMERACY AND RESPONSE STYLE

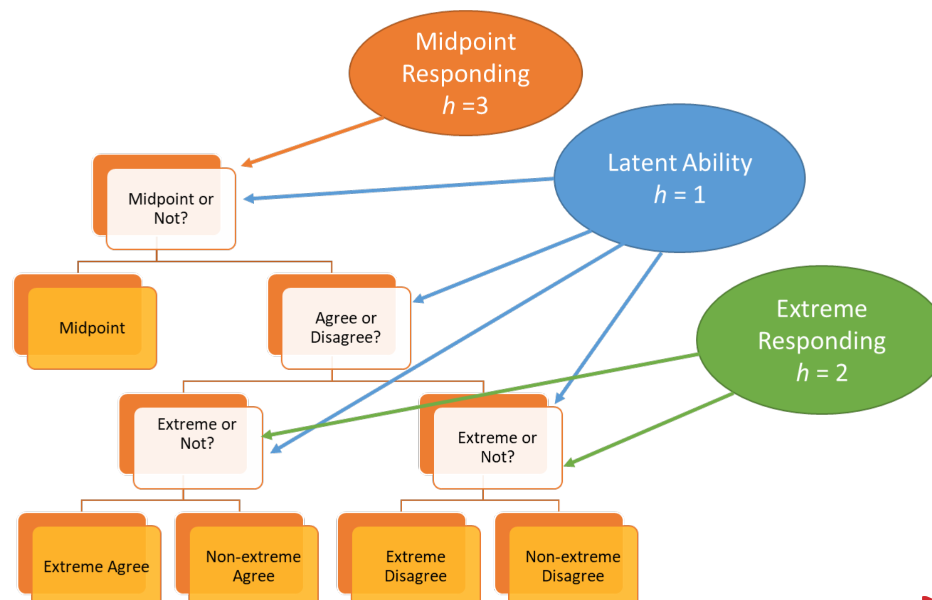
Response style is the tendency to select certain options on a questionnaire regardless of question content.  Extreme and midpoint responding are commonly explored and can add to measurement error and bias estimates from statistical models. Numeracy is a person's understanding of numbers.


**Goal:** Evaluate whether individuals with low numeracy also tend to use extreme and midpoint responding.

**Methods:** IRTree models are used to partial out extreme and midpoint responding based on questionnaire data from the Understanding America Study.

**Results:** Numeracy is negatively related to both extreme and midpoint responding. Even after controlling for education.

**Conclusion:** Studies evaluating numeracy's relationship to a outcome measured on a self-report scale need to be cautious. The Subjective Numeracy Scale is potentially risky as it conflates response style and numeracy.



# POWER FOR WHAT?

**Work with SPSP 2019 Working Group on Power**

Manuscript providing a general overview of practices in power and sample size planning.

- Addresses common misunderstandings about power/power analysis
- Centralize tools for power analysis in a variety of analyses
- Advocates for effect size sensitivity analysis, as sample size is frequently limited by other resources.

Giner-Sorolla, **Montoya**, Carpenter, Lewis, Jr., Aberson, Bostyn, Conrique, Ng, Reifman, Schoemann, Soderberg (under review). https://osf.io/jnmya/

# POWER ANALYSIS IN MEDIATION

Previous research suggests mediation analyses are often under powered.

We conducted a literature review of manuscript published in *Psychological Science* that used mediation analysis:

- No papers provided a justification for sample size related to mediation

Existing tools for power analysis in mediation all use different inferential methods.

We explored whether inferential method used for sample size determination affects the power if a percentile bootstrap CI is used in the data analysis.

**Joint significance test** can be used in lieu of bootstrapping in power analysis, providing very similar sample size recommendations.

**Benefits:** Computational ease

Fossum & Montoya (under review)

# REPLISIMS

**Large scale replication attempt of high impact simulation studies**

Our lab contributed 2 of 8 replications:

- Fritz & MacKinnon (2007): Jessica Fossum
- MacKinnon et al., (2004): Tristan Tibbe

Identified areas to improve….

- Reporting of methods
- Reporting of results
- Open Code/Open Data

Lohmann, Luijken, Alter, Claramunt Gonzalez, Clouth, Fossum, Hesen, Huizing, Ketelaar, **Montoya**, Nab, Nijman, Penning de Vries, Tibbe, Wang, Groenwold (under revision)

# WEIGHT STIGMA STUDY

**Event contingent ecological momentary assessment**, participants report when they experience a weight stigmatizing event for 2 weeks.

> Complete a cortisol sample at time of reporting, 1 hour later, and 25 hours later

**Primary Outcomes:** Cortisol and food consumption

**Analyses:** Multilevel (Poisson) models to examine effects of event and moderators of these effects.

**Results:** We find that men's eating is more susceptible to these events

Hunger, **Montoya**, Hubbard, Edrosolan, Tan, Tomiyama, (in prep). Ecological Momentary Assessment of Weight Stigma and Eating Behavior in Everyday Life.

# JOURNAL EDITORS DISCUSSION INTERFACE

**A guide to adopting open science practices for social science journal editors**

Participated in a hack-a-thon at SIPS 2022, and contributed to drafting and editing of the manuscript.

Approaching open science as a buffet of options

We discuss:

- What it is?
- How to implement it?
- What are the potential roadblocks?

For a number of open science practices. My focus was primarily on Registered Reports.

# JOURNAL CENSUS

We identified 278 journals over 3 waves of data collection

**Research Goals:**

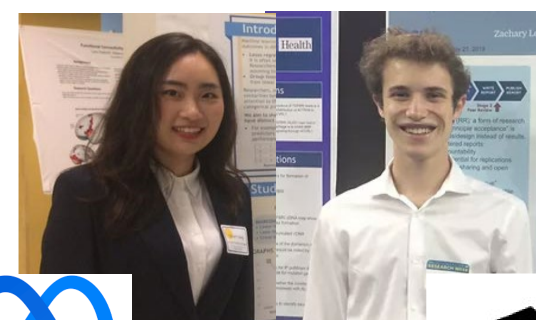
1) Document initial journal adoption

2) Evaluate commonly generated barriers, based on workshops

**Variables (25):**

Registered replications, exploratory data analysis, withdraw after stage 1 review, word limits, author-blind review, etc.

Two independent coders double coded every journal.




Montoya, Krenzer, Fossum, 2021

# REGISTERED REPORTS VS. TRADITIONAL PAPERS

**Conduct survey study with authors of RRs and traditional papers to:**

1) Evaluate if QRPs/OSPs differ between authors of RRs and traditional papers

2) Evaluate if authors who published a RR consider themselves less likely to use QRPs/more likely to use OSPs in the future

3) Compare RRs and traditional papers on time to publication

4) Compare the impact of RRs and traditional papers (using citation counts)

**Matched-pairs design**
Controls for publication timing, journal quality, peer review process, type of researcher


Montoya, Tibbe, Krenzer, *In-principle Acceptance, RSOS*

# INTERLEAVING GENERALIZABILITY

**Large scale educational study of the impact of interleaved homework assignments on learning in STEM**

Interleaving has been demonstrated to benefit learning in STEM, but not evaluated on a large scale. We implement an experimental design in large college classes in a variety of academic fields (psychology, math, physics, biology, engineering).

**Goal:** Evaluate the generalizability of the interleaving benefit and estimate cross-field and cross-classroom heterogeneity in this effect.

**Methods**: Multi-class within-class designs, half of topics are interleaved and half are blocked. Performance on exam items used as primary outcome.

**Analysis**: 3-level multilevel models used to estimate heterogeneity

Data collection begins January 2023

# SAMPLE SIZE PLANNING IN MULTILEVEL MODELING

Two papers by Maas & Hox (2004, 2005) are some of the most highly cited for sample size justification in multilevel modeling.
The simulation is limited in scope so we conducted a large scale literature review to examine the context in which these papers are cited.

Replication of original study and extension based on literature search

- Unbalanced clusters
- Correlations among random effects
- Lower ICC
- Varied number of predictors and interactions

Evaluate whether general recommendation from Maas & Hox holds in broader conditions

Generate reporting guidelines (based on lit search)

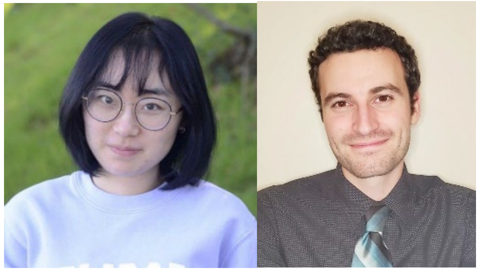
Sample size planning recommendations

**Being submitted as a registered report**

Guassi Moreira, Block, Cruz, Enders, Montoya (in prep)

# LASSO REGRESSION WITH MULTICATEGORICAL PREDICTORS

Linear regression can fit models with multicategorical predictors represented multiple ways without effecting fit of the model, but it does change the coefficients (because their interpretation changes)



**Goal:** Evaluate whether Lasso and Group lasso are sensitive to coding strategy because they are regularization methods and penalize based on the size of coefficients.

**Methods:** Large scale data analysis and Monte Carlo simulation

**Results:** Lasso and group lasso's prediction accuracy are both affected by coding strategy. Variable selection is only affected for lasso. We developed an algorithm which seeks to find the coding strategy with best prediction accuracy.

Huang, Tibbe & Montoya (in prep). https://psyarxiv.com/wc45u/

# DETECTING TREATMENT RESPONDERS IN RCTS

There are multiple AUD medications available, and we might expect that some individuals with respond to specific mediation.

Using data from 4 RCTs we evaluate machine learning methods to identify these sets of treatment responders.

- Qualitative Interaction Trees (QUINT)
  - Prioritizes qualitative interactions
  - Difficult to interpret complex trees
- Group Lasso Interaction Nets (Glinternet)
  - Does not prioritize qualitative interactions
  - Easy to interpret

We provide recommendations for which methods to use and how in the context of detecting treatment responders using RCT data.

Grodin, Montoya, Cruz, Donato, Baskerville, Ray, (under revision) https://psyarxiv.com/uvn3m/