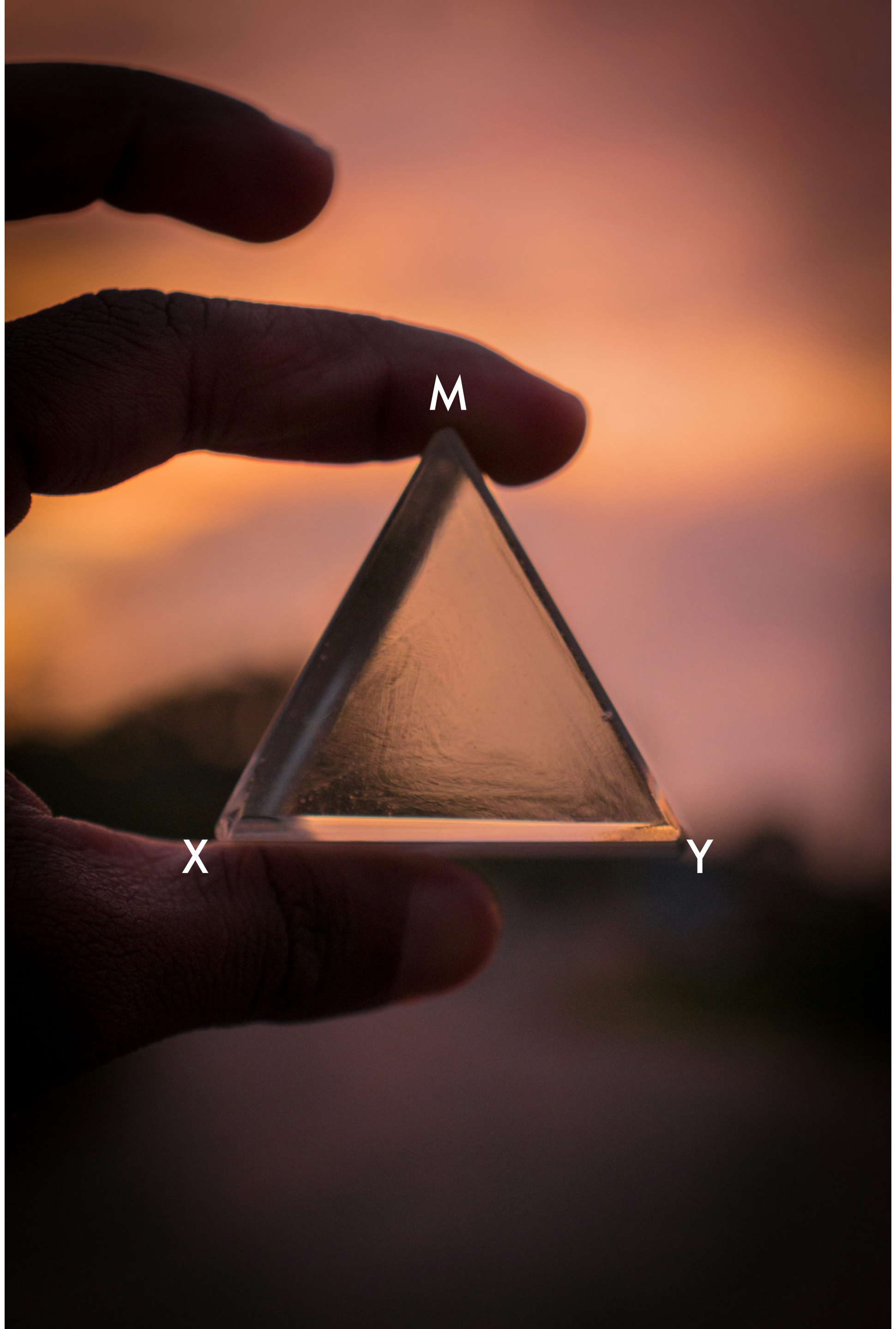


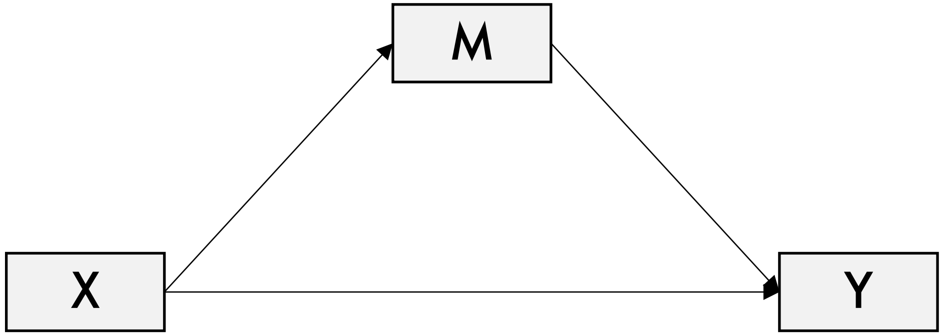
Trust the process?
(Causal) Mediation analysis

Julia M. Rohrer
Leipzig University
juliarohrer.com
www.the100.ci

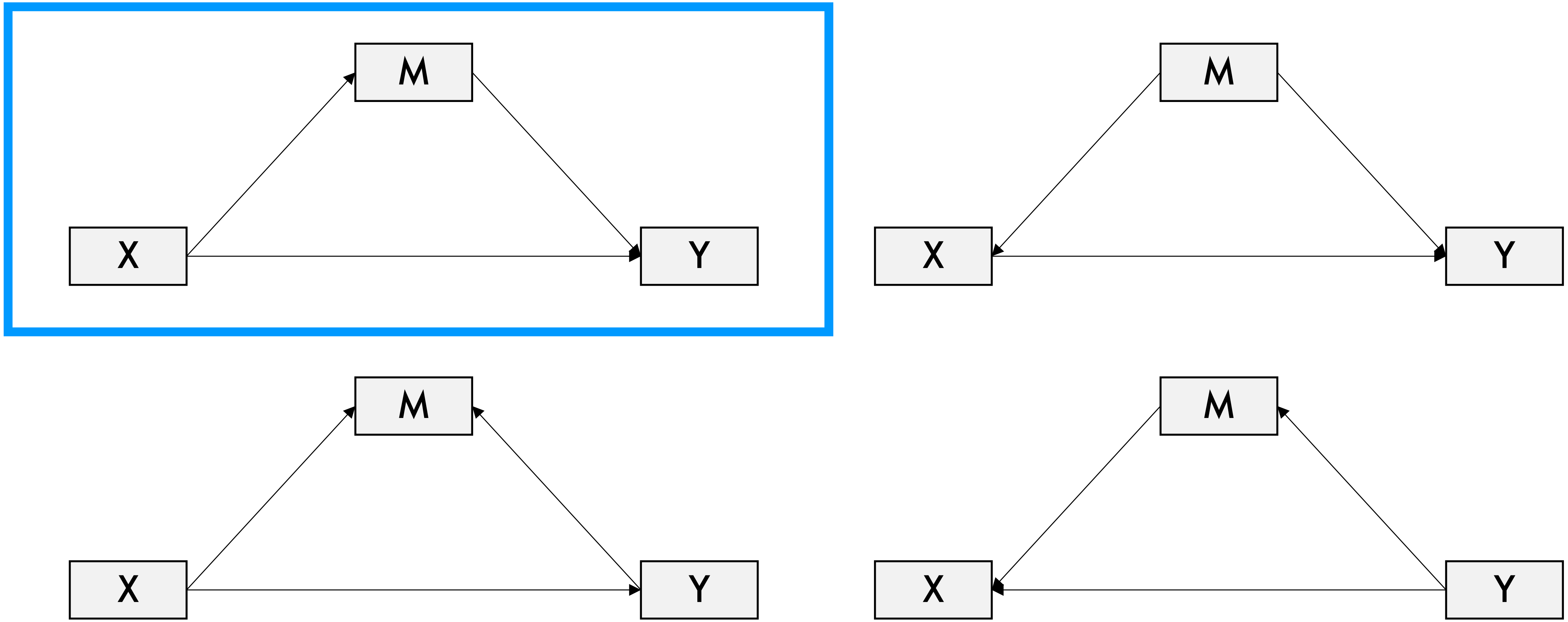
Slides: Resources

Foto von Apurv Das auf Unsplash

Mediation analysis as a causal inference problem



- Two central questions
  - To which extent does X affect Y via M? (Indirect effect)
  - To which extent does X affect Y apart from its effect via M? (direct effect)
- This is unambiguously a *causal* research question
  - Baron & Kenny, 1986; MacKinnon, 2008; Preacher, 2015



In all of these scenarios we could fit a mediation model with M as a mediator of the effect of X on Y and find significant „direct and indirect effects.“

Only in one of the scenarios could the estimates possibly have the desired interpretations (direct and indirect effects).

(If you don‘t desire that interpretation, you‘re not doing a mediation analysis but are instead for some reason investigating conditional associations)

Three steps of mediation analysis

(or really any causal analysis), Nguyen et al. (2020)

1. Definition
   - » Define the causal effects of interest
2. Identification
   - » Figure out the assumptions under which the effects of interest can actually be identified from the data (assuming that an infinite data were available)
3. Estimation
   - » Actually estimate the effects of interest (using only the available finite data)

Three steps of mediation analysis

(or really any causal analysis), Nguyen et al. (2020)

In the „traditional" mediation approach, this is a one-step procedure with a focus on estimation:

1. The effects of interest are defined by the parameters of the statistical model.
2. The identification assumption is essentially that the statistical model matches the actual model underlying reality.
3. Estimation via e.g. Baron & Kenny's (1986) three steps (not optimal), SEM (better), the regression-based approach implemented in PROCESS

Three steps of mediation analysis

(or really any causal analysis), Nguyen et al. (2020)

In the „traditional" mediation approach, this is a one-step procedure with a focus on estimation:

1. The effects of interest are defined by the parameters of the statistical model.
2. The identification assumption is essentially that the statistical model matches the actual model underlying reality.
3. Estimation via e.g. Baron & Kenny's (1986) three steps (not optimal), SEM (better), the regression-based approach implemented in PROCESS

The central point of contention

- Mediation analysis rests on very strong identification assumptions – which are likely to be wrong in many applications in psychology
  - Bullock et al. (2010): Yes, but what's the mechanism? (Don't expect an easy answer)
  - Fiedler et al. (2011): What mediation analysis can (not) do
  - Rohrer et al. (2022): That's a lot to process! Pitfalls of popular paths models
  - Blog posts: In psychology, everything mediates everything; Indirect Effect Ex Machina; Mediation analysis is counterintuitively invalid; That's a very nice mediation analysis you have there. It would be a shame if something happened to it.

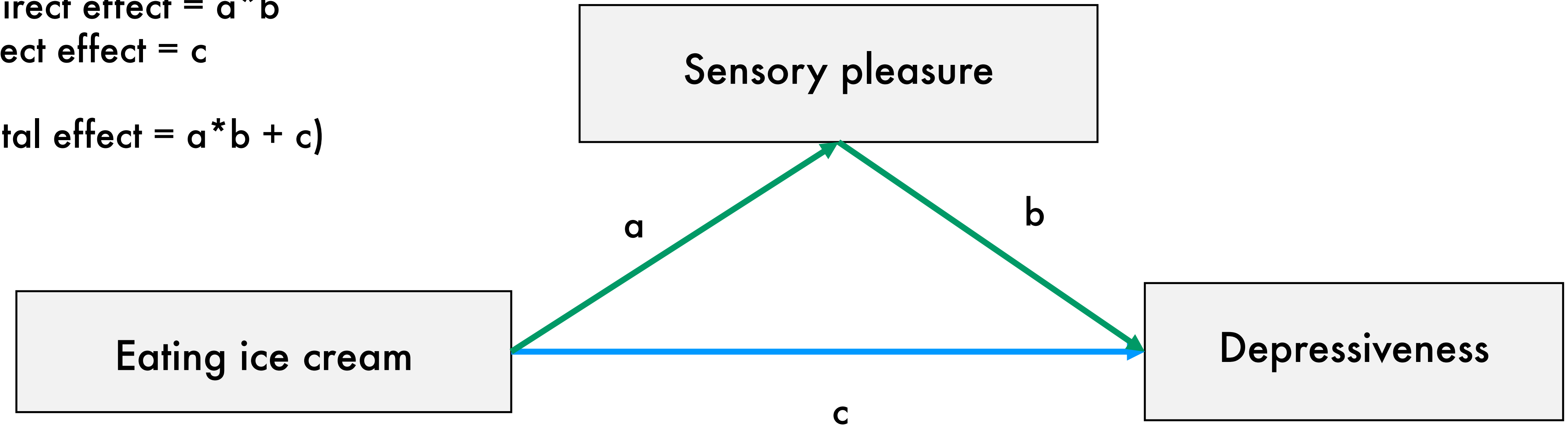
2. Identification:

We need to be able to estimate the causal effects a, b, and c *without any bias.*

1. Definition of the causal effects:

Indirect effect = a*b
Direct effect = c

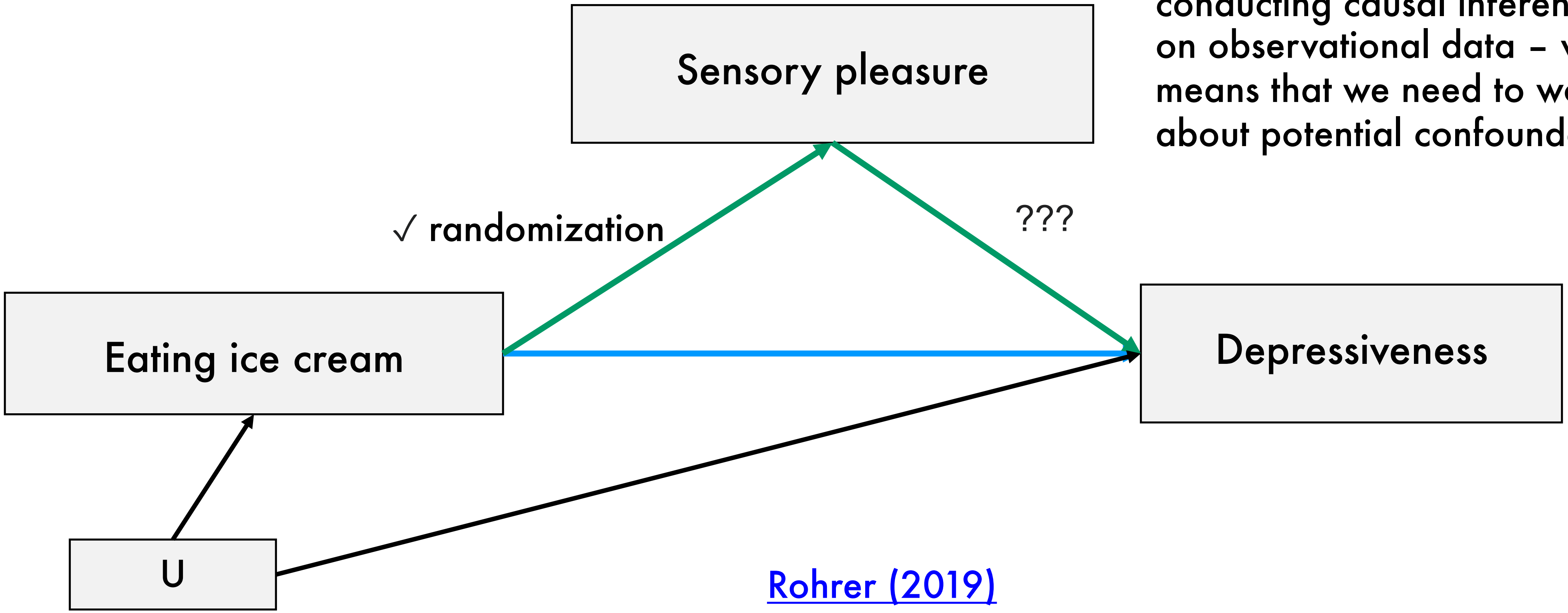
(Total effect = a*b + c)



Rohrer (2019)

Sensory pleasure has not been randomized.

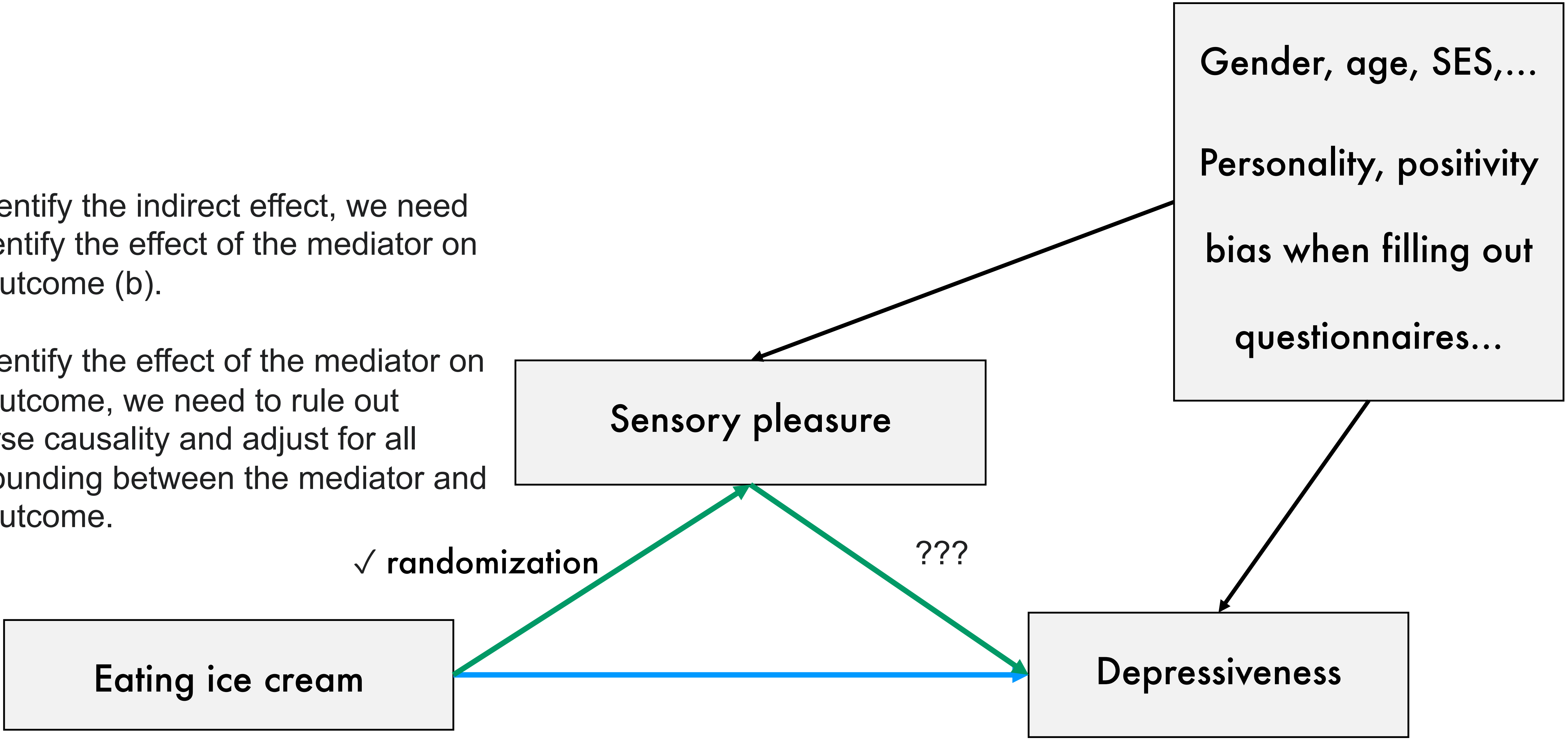
Thus, when we try to identify the b-path (Sensory Pleasure → Depressiveness), we are conducting causal inference based on observational data – which means that we need to worry about potential confounders.



Rohrer (2019)

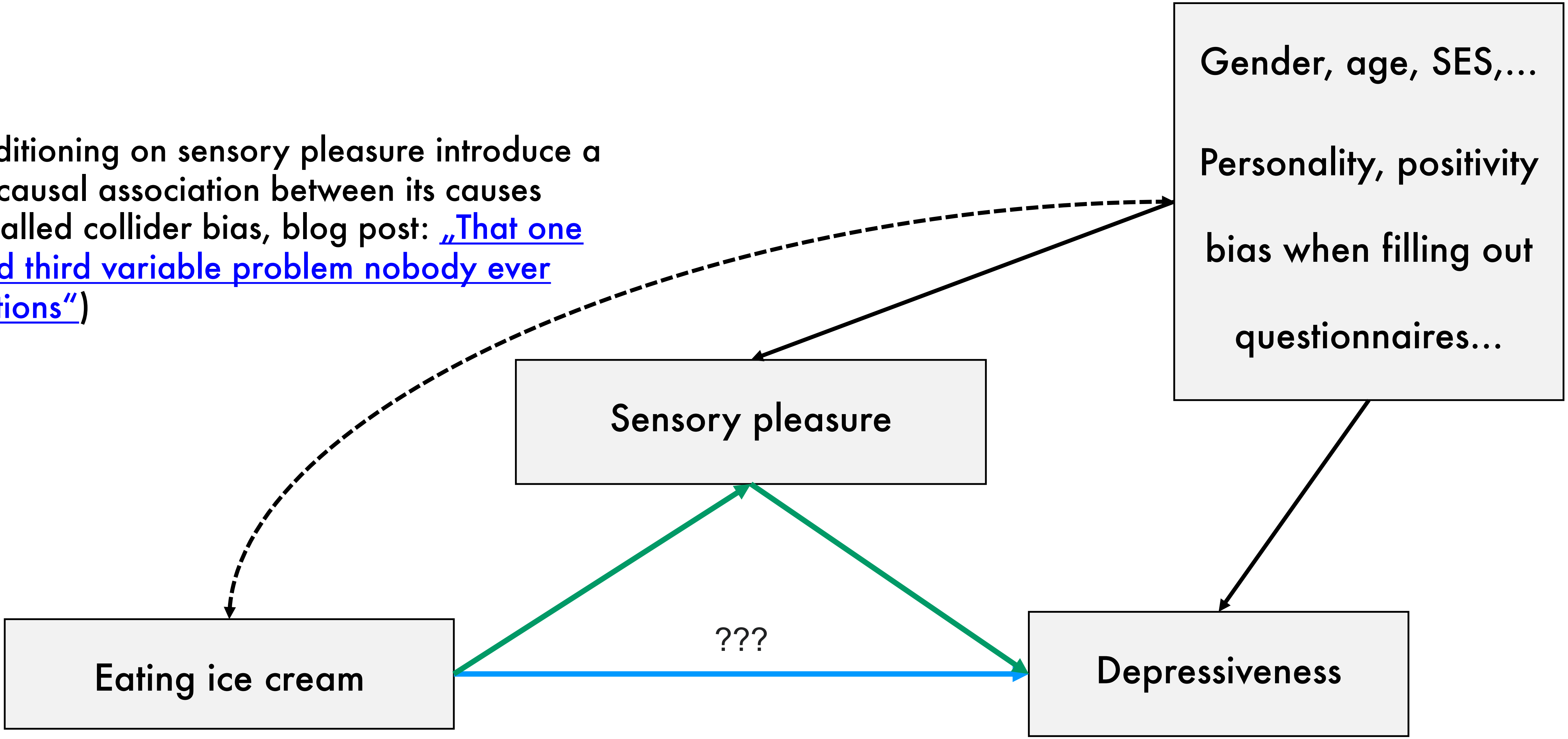
To identify the indirect effect, we need to identify the effect of the mediator on the outcome (b).

To identify the effect of the mediator on the outcome, we need to rule out reverse causality and adjust for all confounding between the mediator and the outcome.



Rohrer (2019)

Conditioning on sensory pleasure introduce a non-causal association between its causes (so-called collider bias, blog post: „That one weird third variable problem nobody ever mentions“)



Gender, age, SES,...

Personality, positivity bias when filling out questionnaires...

Sensory pleasure

???

Eating ice cream

Depressiveness

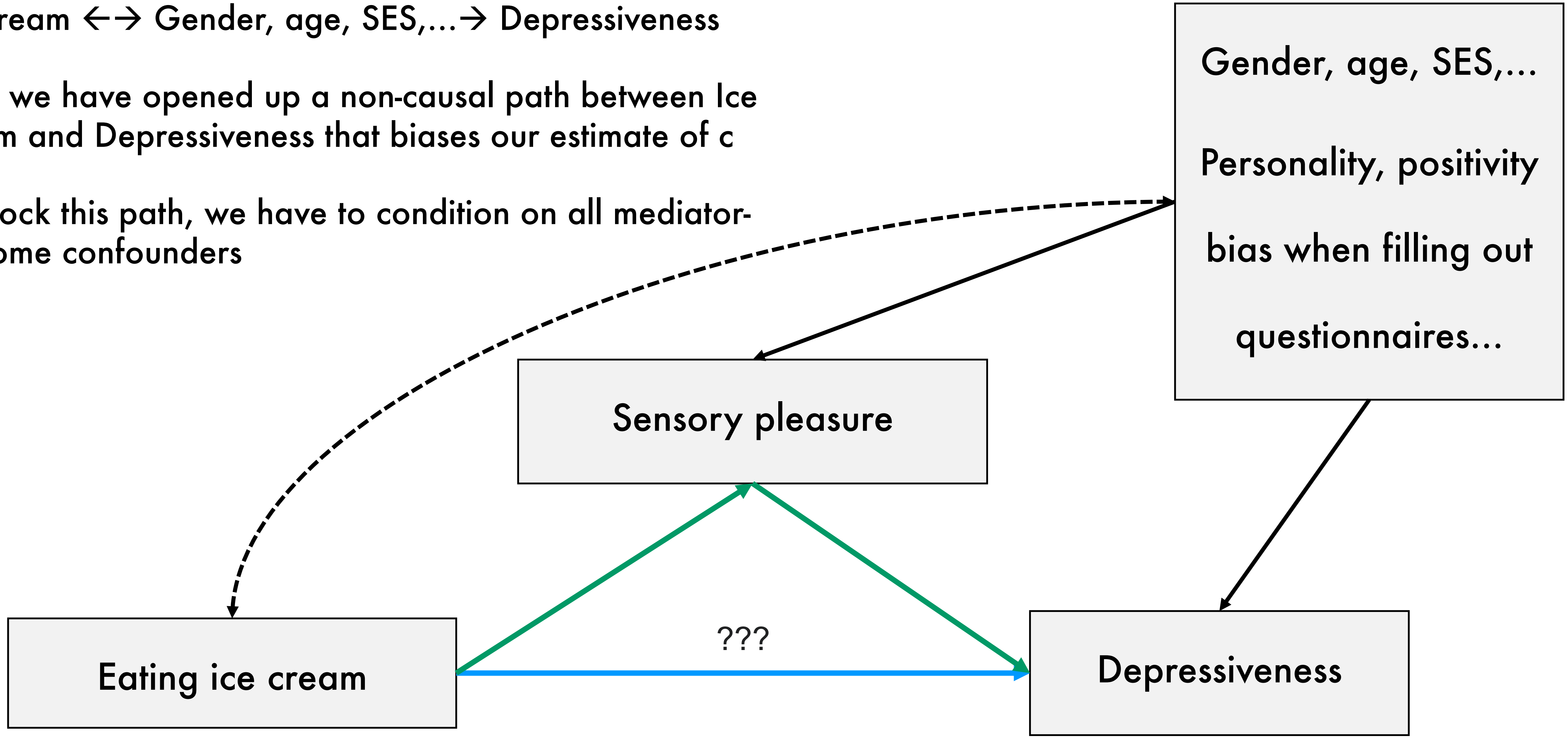
Ice Cream -> Sensory pleasure -> Depressiveness; block the indirect effect by conditioning on sensory pleasure

Ice Cream -> Sensory pleasure <- Positivity bias etc. -> Depressiveness

Ice cream ←→ Gender, age, SES,...→ Depressiveness

Now we have opened up a non-causal path between Ice cream and Depressiveness that biases our estimate of c

To block this path, we have to condition on all mediator-outcome confounders



Gender, age, SES,...

Personality, positivity bias when filling out questionnaires...

Sensory pleasure

Eating ice cream

Depressiveness

Ice Cream -> Sensory pleasure -> Depressiveness; block the indirect effect by conditioning on sensory pleasure

Ice Cream -> Sensory pleasure <- Positivity bias etc. -> Depressiveness

Identification assumptions of mediation analysis

- In a nutshell, to identify both the indirect and the direct effect, we have to
  - *know* all mediator-outcome confounders
  - have measured them appropriately and
  - have statistically adjusted for them appropriately (e.g., by including them as a covariate in our regression analyses, by including them in our SEM)
- If the cause of interest (X) was not randomized, the same applies to any confounders between X and Y
  - And X and M (these are a subset of the confounders between X and Y)

Identification assumptions of mediation analysis

- These assumptions are usually not realistic in applications in psychology
  - Both mediator and outcome are usually psychological variables which are potentially confounded by a myriad of other factors
- One weird trick (Rohrer, 2019)
  - Pick a mediator that is
    - plausibly affected by X (X → M)
    - plausibly *confounded* with Y (M ← C → Y)
  - You now have good chances to get a significant (but spurious) indirect effect of X on Y via M, even if X does not affect Y at all

Identification assumptions of mediation analysis

- The assumptions about no unobserved confounding apply to both the traditional approach and more modern „causal“ appraoches
- Additional assumptions in the traditional approach (indirect effect = a*b)
  - Effect of M on Y does not depend on X (no treatment-mediator interaction; MacKinnon et al., 2020)
  - The effects X → M and M → Y
    - Are the same for everyone or alternatively
    - May vary, but do so independently (otherwise, the multiplicative logic does not work out)

These may also often be implausible, in which case a more modern „causal“ approach may make sense

„Causal“ mediation analysis

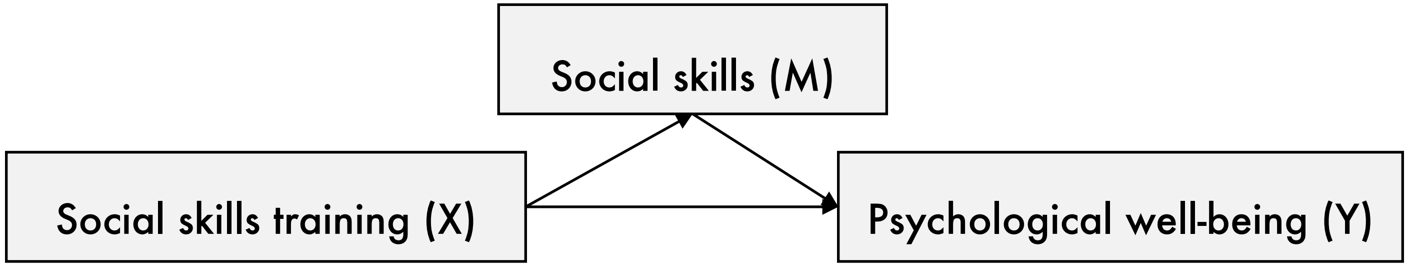
- *Causal* because they are explicitly (and actively) being developed in the causal inference literature (e.g., Imai et al., 2010)
  - The analysis goals are precisely as causal as the analysis goals of traditional mediation analysis
  - All problems concerning potentially unobserved confounding still very much apply

„Causal“ mediation analysis: Three steps

- More focus on the precise definition of the effects of interest
- Usually more awareness of the necessary identifiation assumptions
  - (lack of unobserved confounding is usually assumed)
- More flexible and varied estimation approaches
  - Allowing e.g. for non-linear effects, treatment-mediator interactions etc.
  - Implemented in various *R* packages

More precise definitions of (in)direct effects

Nguyen et al. (2020)



- » Counterfactuals in the individual-level (total) causal effects of X are simple
  - » Effect of X on Y *for you*: Your psychological well-being if you participated in the training minus your psychological well-being if you didn‘t participate in the training
  - » In potential outcomes notation: $Y^{X=1} - Y^{X=0}$
- » But which potential outcomes are we considering when talking about direct and indirect effects?
  - » This will involve cross-world counterfactuals – e.g., what if you received the training but your social skills did not change (e.g., what if the mediator remained constant?)

»Total effect $= Y^{X=1} - Y^{X=0} = \underbrace{Y^{X=1,\ M(X=1)}} - \underbrace{Y^{X=0,\ M(X=0)}}$

Your outcome
if you receive treatment
and
the mediator takes on the value it takes on when you receive treatment

Your outcome
if you **do not** receive treatment
and
the mediator takes on the value it takes on when you **do not** receive treatment

(e.g., your well-being after the skills training if you have the social skills you'd naturally acquire from the training)

» Total effect = $Y^{X=1} - Y^{X=0} = Y^{X=1,\ M(X = 1)} - Y^{X=0,\ M(X = 0)}$

» Natural direct effect = $\underbrace{Y^{X=1,\ M(X = 0)}} - \underbrace{Y^{X=0,\ M(X = 0)}}$

Your outcome
if you receive treatment
and
the mediator takes on the value it takes on when you **do not** receive treatment

Your outcome
if you **do not** receive treatment
and
the mediator takes on the value it takes on when you **do not** receive treatment

(e.g., your well-being after the skills training if your social skills remain on the level if you did not receive the training)

Contrast gives essentially the effect of „just training, no skills acquired"

- Total effect $= Y^{X=1} - Y^{X=0} = Y^{X=1,\, M(X=1)} - Y^{X=0,\, M(X=0)}$
- Natural direct effect $= Y^{X=1,\, M(X=0)} - Y^{X=0,\, M(X=0)}$
- Natural indirect effect $= Y^{X=1,\, M(X=1)} - Y^{X=1,\, M(X=0)}$

Your outcome
if you receive treatment
and
the mediator takes on the value it takes on when you **do** receive treatment

Your outcome
if you **do** receive treatment
and
the mediator takes on the value it takes on when you **do not** receive treatment

Contrast gives essentially the effect of „skills on top of training“

- Total effect = $Y^{X=1} - Y^{X=0} = Y^{X=1,\ M(X = 1)} - Y^{X=0,\ M(X = 0)}$
- Natural direct effect = $Y^{X=1,\ M(X = 0)} - Y^{X=0,\ M(X = 0)}$
- Natural indirect effect = $Y^{X=1,\ M(X = 1)} - Y^{X=1,\ M(X = 0)}$
- This is the *direct-indirect* composition
  - We start from the control condition $Y^{X=0,\ M(X = 0)}$ and first add the direct effect to end up at $Y^{X=1,\ M(X = 0)}$, then add the indirect effect to end up in the treatment condition $Y^{X=1,\ M(X = 1)}$

- Alternatively we could do an indirect-direct composition and first add the indirect effect, then the direct effect
  - Natural indirect effect = $Y^{X=0,\ M(X=0)} - Y^{X=0,\ M(X=1)}$
  - Natural direct effect = $Y^{X=0,\ M(X=1)} - Y^{X=1,\ M(X=1)}$
- The direct-indirect and indirect-direct composition will return different direct and indirect effects *whenever there is a treatment-mediator interaction*
  - Essential question: does that interaction count toward the direct effect or toward the indirect effect?
    - VanderWeele (2014) offers a unified decomposition into 4 components that keeps the interaction separate

- Which decomposition is the right one?
  - „Is there a mediated effect on top of whatever direct effect there is?“: direct-indirect decomposition
  - „Is there a remaining direct effect on top of whatever indirect effect there is?“: indirect-direct decomposition
  - When in doubt, both
- In either case, we usually can only aim to estimate the *averages* of these individual-level direct and indirect effects

Less natural effects that can be estimated

- Interventional (in)direct effects: Contrasting hypothetical interventions that fix treatment or mediator at specific values
  - These do *not* decompose the total effect, but rather answer specific targeted research questions
  - E.g., Sexual minority status → Bullying experiences → Depressive symptoms
    - „What would be the effect of sexual minority status on depressive symptoms if we intervened so that the experience of bullying experiences among sexual minorities is the same as bullying experiences among majority members?“
  - Controlled direct effects: Effect of X if we hold M constant at a certain level, e.g.: Effect of sexual minority status when we hold bullying experiences constant at „no bullying at all“

Three steps of mediation analysis

(or really any causal analysis), Nguyen et al. (2020)

1. Definition
   - » Define the causal effects of interest
2. Identification
   - » Figure out the assumptions under which the effects of interest can actually be identified from the data (assuming that an infinite data were available)
3. Estimation
   - » Actually estimate the effects of interest (using only the available finite data)

Identification assumptions in causal mediation analysis

- Unfortunately, we still have to assume that all confounding has been adjusted appropriately
- The specific assumptions depend on the effects of interest
    - Total effect obviously the easiest one (randomization does the trick)
    - Interventional effects (e.g., controlled direct effects) are harder
        - Have to assume no unobserved mediator-outcome confounding
    - Natural direct and indirect effects are the hardest
        - Intermediate confounders are not allowed to exist

Three steps of mediation analysis

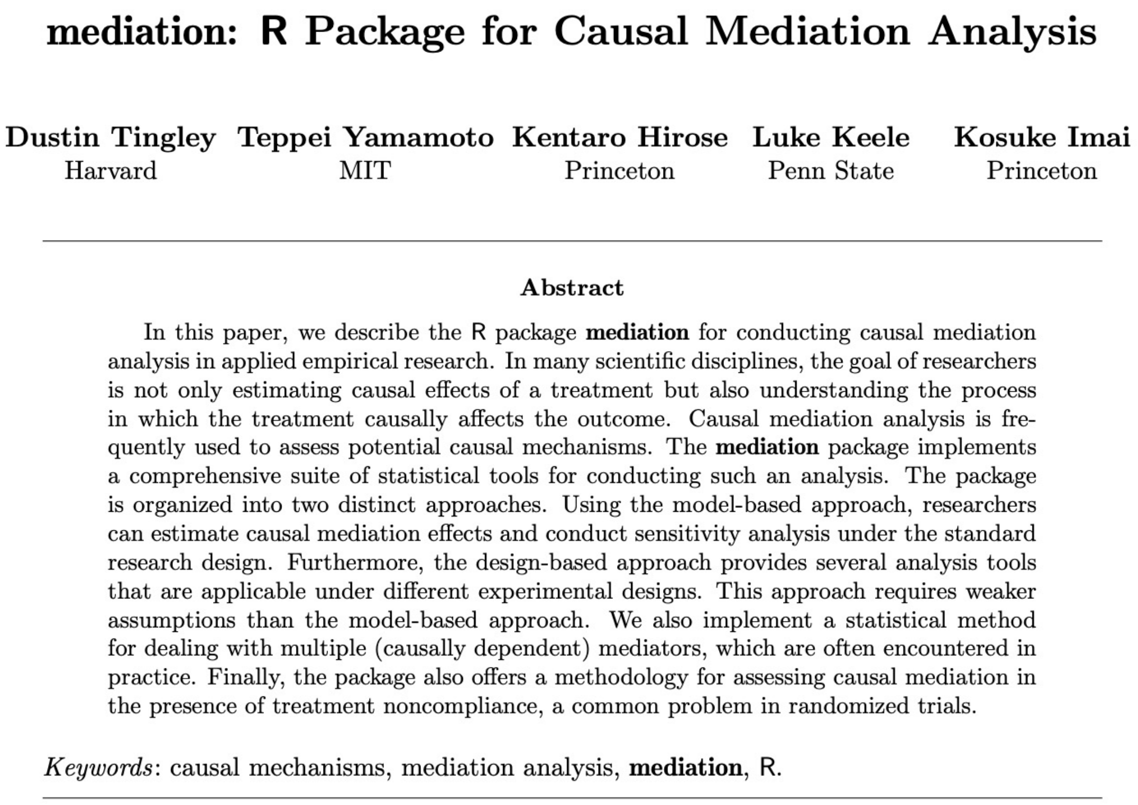
(or really any causal analysis), Nguyen et al. (2020)

1. Definition
   - » Define the causal effects of interest
2. Identification
   - » Figure out the assumptions under which the effects of interest can actually be identified from the data (assuming that an infinite data were available)
3. Estimation
   - » Actually estimate the effects of interest (using only the available finite data)

Estimation of causal mediation

- This is essentially just an extension of how causal effects are estimated more generally
  - Often combination of multiple models (model for the outcome mean Y, model for the mediator distribution M, model for the exposure assignment X)
- Common approaches, both parametric and non-parametric
  - Regression
  - Weighting
  - „Imputation“, g-computation (Wang & Arah, 2015)
  - See Chatton & Rohrer (2024) for a conceptual introduction to these approaches outside of the mediation context

Software!



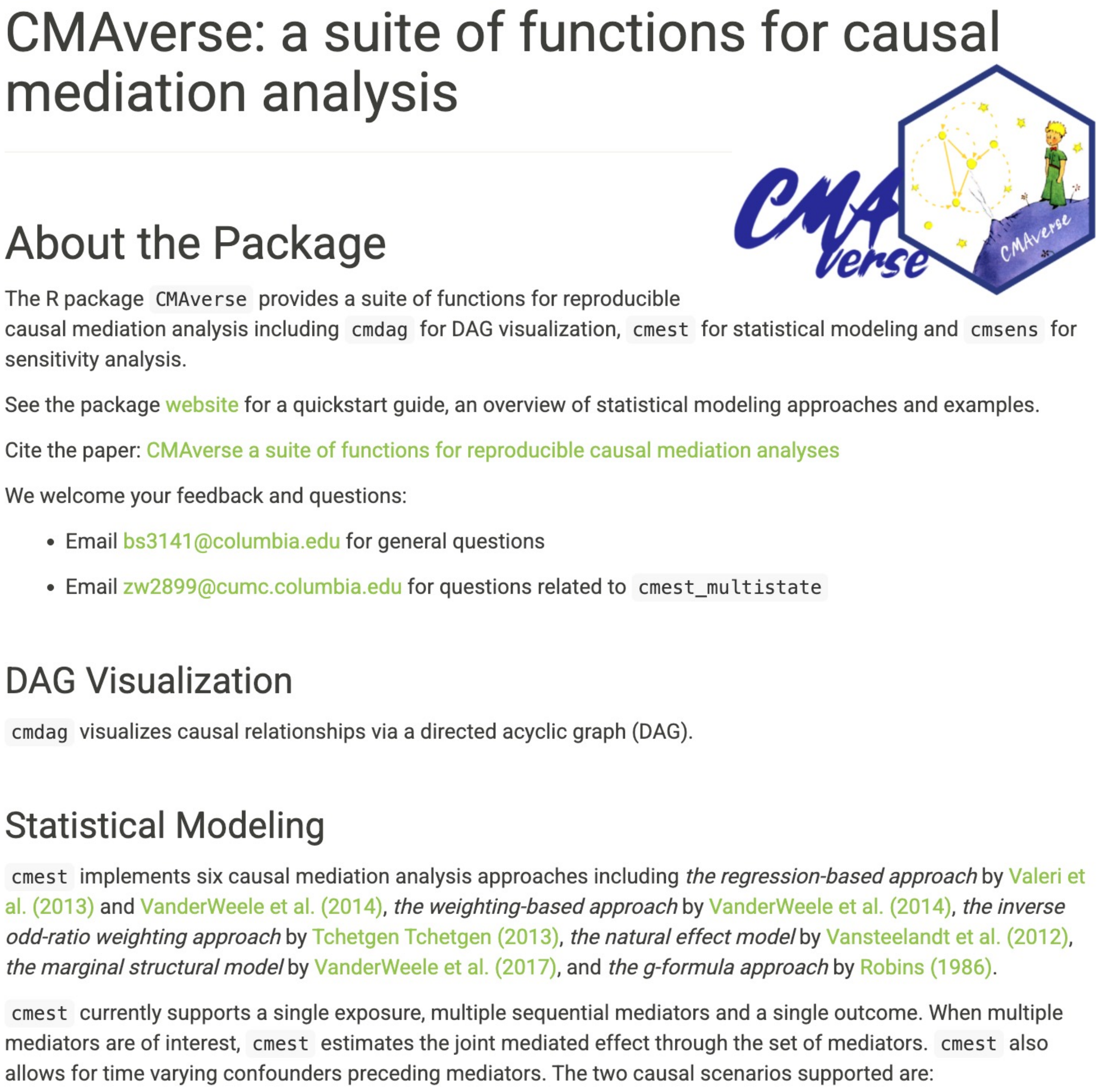
mediation: R Package for Causal Mediation Analysis

Dustin Tingley, Harvard
Teppei Yamamoto, MIT
Kentaro Hirose, Princeton
Luke Keele, Penn State
Kosuke Imai, Princeton

Abstract

In this paper, we describe the R package **mediation** for conducting causal mediation analysis in applied empirical research. In many scientific disciplines, the goal of researchers is not only estimating causal effects of a treatment but also understanding the process in which the treatment causally affects the outcome. Causal mediation analysis is frequently used to assess potential causal mechanisms. The **mediation** package implements a comprehensive suite of statistical tools for conducting such an analysis. The package is organized into two distinct approaches. Using the model-based approach, researchers can estimate causal mediation effects and conduct sensitivity analysis under the standard research design. Furthermore, the design-based approach provides several analysis tools that are applicable under different experimental designs. This approach requires weaker assumptions than the model-based approach. We also implement a statistical method for dealing with multiple (causally dependent) mediators, which are often encountered in practice. Finally, the package also offers a methodology for assessing causal mediation in the presence of treatment noncompliance, a common problem in randomized trials.

Keywords: causal mechanisms, mediation analysis, **mediation**, R.



CMAverse: a suite of functions for causal mediation analysis

About the Package

The R package `CMAverse` provides a suite of functions for reproducible causal mediation analysis including `cmdag` for DAG visualization, `cmest` for statistical modeling and `cmsens` for sensitivity analysis.

See the package website for a quickstart guide, an overview of statistical modeling approaches and examples.

Cite the paper: CMAverse a suite of functions for reproducible causal mediation analyses

We welcome your feedback and questions:

- Email bs3141@columbia.edu for general questions
- Email zw2899@cumc.columbia.edu for questions related to `cmest_multistate`

DAG Visualization

`cmdag` visualizes causal relationships via a directed acyclic graph (DAG).

Statistical Modeling

`cmest` implements six causal mediation analysis approaches including *the regression-based approach* by Valeri et al. (2013) and VanderWeele et al. (2014), *the weighting-based approach* by VanderWeele et al. (2014), *the inverse odd-ratio weighting approach* by Tchetgen Tchetgen (2013), *the natural effect model* by Vansteelandt et al. (2012), *the marginal structural model* by VanderWeele et al. (2017), and *the g-formula approach* by Robins (1986).

`cmest` currently supports a single exposure, multiple sequential mediators and a single outcome. When multiple mediators are of interest, `cmest` estimates the joint mediated effect through the set of mediators. `cmest` also allows for time varying confounders preceding mediators. The two causal scenarios supported are:

What about design based solutions?*

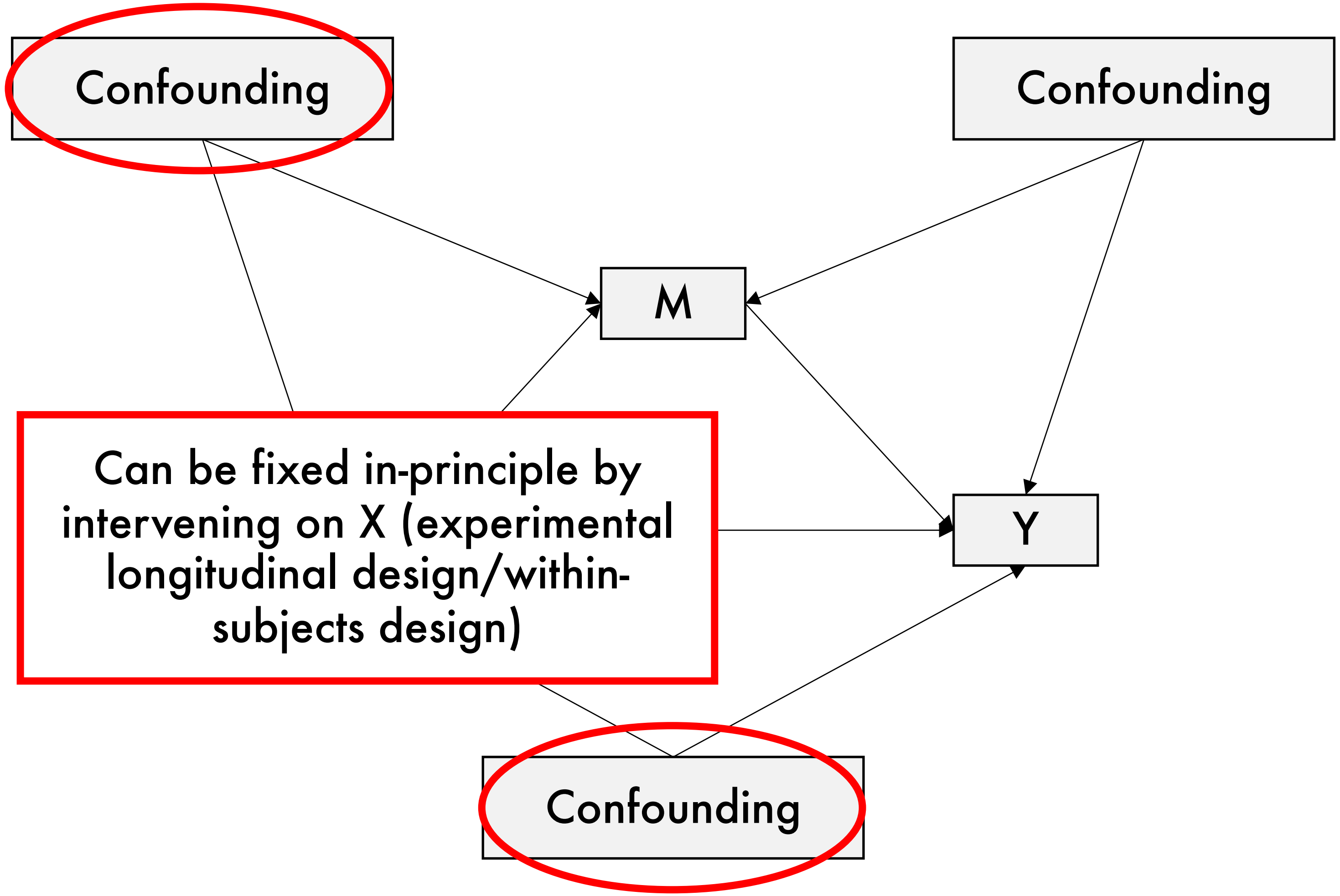
*Disclaimer: I have not kept up with that literature

Design-based solutions

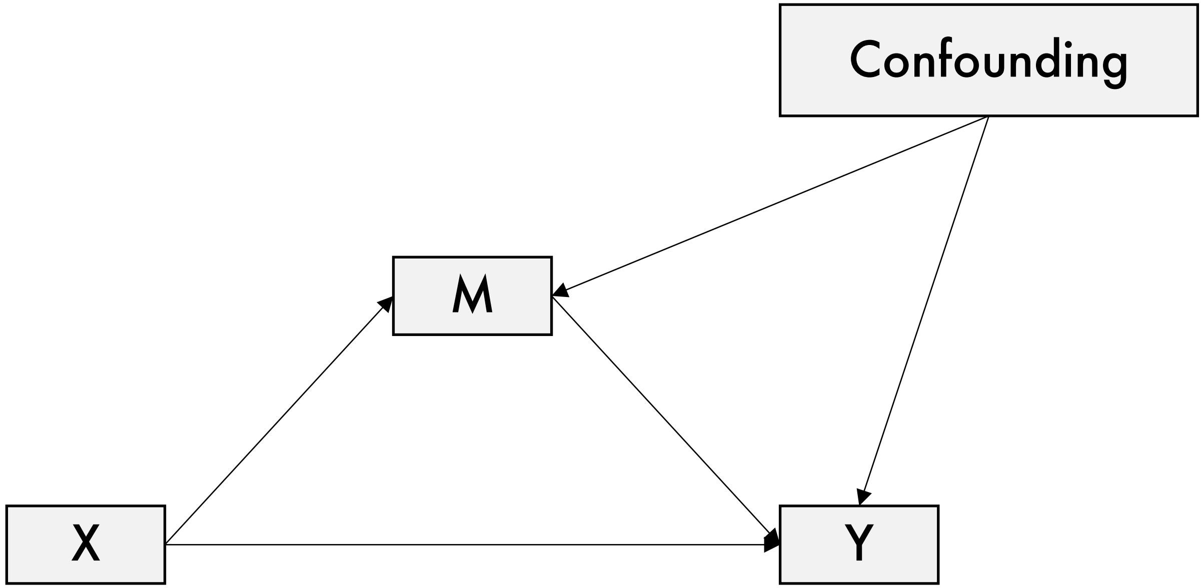
- longitudinal data
- intervening on X *and* on the mediator
    - sequentially
    - or simultaneously

Longitudinal data

- do *not* automatically solve confounding, see Rohrer & Murayama (2023)
- can rule out *certain types* of confounding if analyzed properly



Can be fixed in-principle by intervening on X (experimental longitudinal design/within-subjects design)



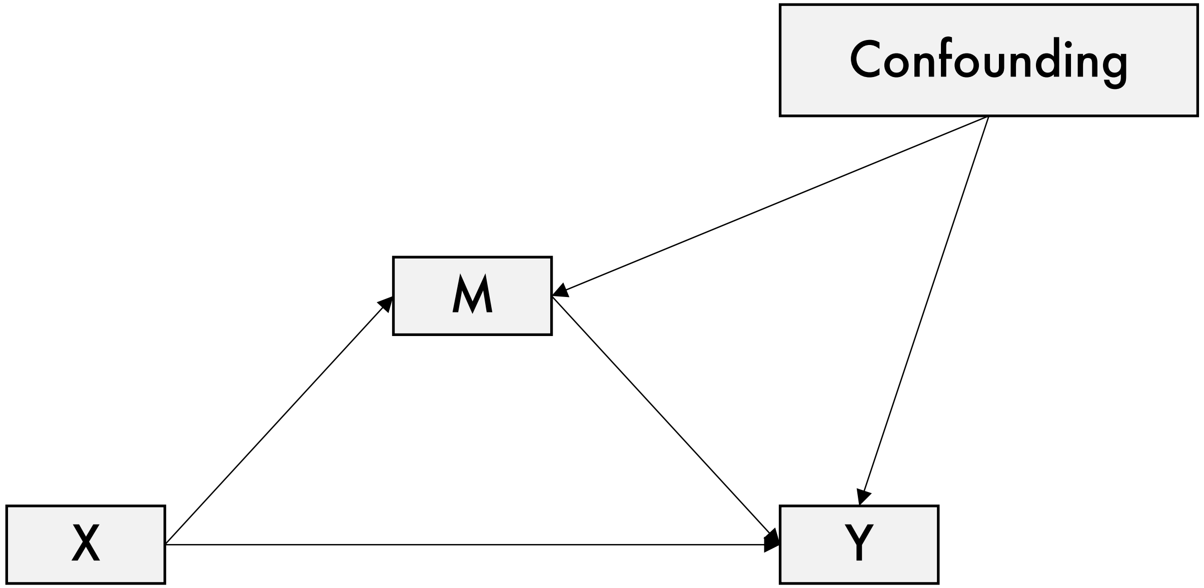
Confounding comes in “two types”

Time-invariant confounding:

- *stable* factors that uniformly affect M and X over the course of the study
- e.g., participants’ gender, socio-economic status, education, stable personality traits and other dispositions…
- what’s time-invariant will depend on the duration of the study (stable over hours vs. days vs. years)

Time-varying confounding:

- fluctuating factors that can change *within* a person, or whose effects may change over time
- e.g., mood, health, any form of psychological state, current living situation



Opportunity afforded by longitudinal data: Compare people with *themselves* over time (within-subject analysis)

→ If we compare people to themselves, they are perfectly "matched" on time-invariant confounders
→ time-invariant confounders cannot "explain away" the observed within-subject associations

Confounding comes in "two types"

Time-invariant confounding:

- *stable* factors that uniformly affect M and X over the course of the study
- e.g., participants' gender, socio-economic status, education, stable personality traits and other dispositions...
- what's time-invariant will depend on the duration of the study (stable over hours vs. days vs. years)

Time-varying confounding:

- fluctuating factors that can change *within* a person, or whose effects may change over time
- e.g., mood, health, any form of psychological state, current living situation

Longitudinal data

- can rule out *time-invariant* confounding
  - if analyzed properly, rules of thumb:
    - in multilevel analyses, you want a fixed effects model or some variation (including person ID as a categorical predictor [not just random intercepts!], within-subject centering, Mundlak device [person-specific mean of the cause of interest as additional predictor])
    - in a SEM context, you can get away with random intercepts *as long as they are allowed to correlate*

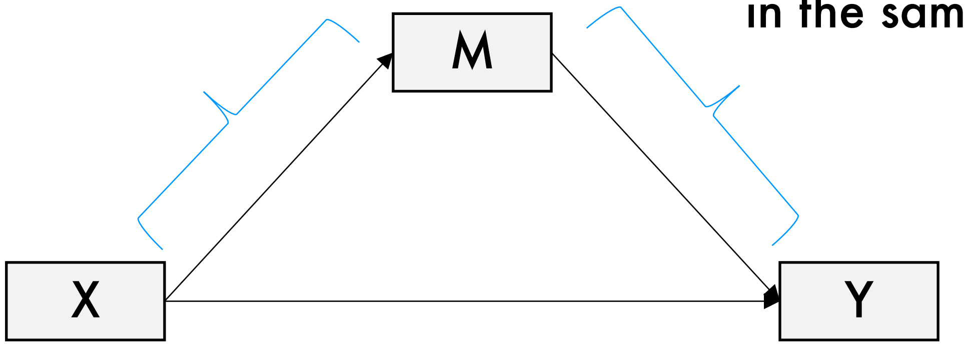
Longitudinal data

- can rule out *time-invariant* confounding
- you still need to worry about time-varying confounding
  - could there be things that change over time within people that confound their M with their Y?
    - If yes, these would need to be additionally controlled for
- "maximal" design-based approach here: repeated within-subject experimentation

Design-based solutions

- additionally intervening on the mediator (either sequentially or simultaneously)
  - *if* that is an option at all – specific problems arise when trying to manipulate for example psychological mediators (Eronen, 2020)
  - assumptions are still needed to chain together the estimated effects, Imai, Keele & Tingley (2010)

Sequential interventions



First experiment identifies the *average* effect of X on M

Second experiment identifies the *average* effect of M on Y, possibly in the same people

- combining the two average effects does *not* necessarily result in the average indirect effect because the individual-level effects could be correlated
  - Scenario 1: X has an effect on M in half of the people; M has an effect on Y in only the other half of people → a and b are non-zero but the indirect effect is zero
  - Scenario 2: X has an effect on M in half of the people; M has an effect on Y in only the same half of people → a and b are non-zero; the indirect effect is *larger* than if one just multiplied the average effects

Person	a	b	a*b
1	1	0	0
2	1	0	0
3	0	1	0
4	0	1	0
Average	0.5	0.5	0

Person	a	b	a*b
1	1	1	1
2	1	1	1
3	0	0	0
4	0	0	0
Average	0.5	0.5	0.5

Sequential interventions

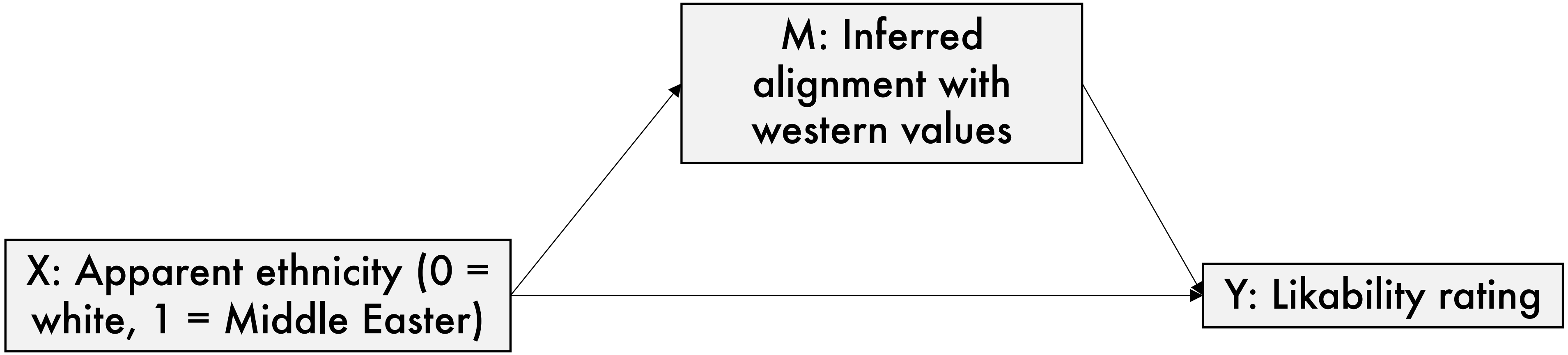
- "maximal" design-based approach: intervene on X and on M *repeatedly* in the same people
  - this way, you can estimate *person-specific* average effects (averaged over time) of X → M and M → Y
  - these can be combined to estimate person-specific average indirect effects
    - necessary assumptions (I think): no time-varying intermediate confounders (mediator-outcome confounders affected by X); person-specific causal homogeneity over the observed window (causal parameters don't change over time)

Simultaneous interventions

- e.g.: Manipulate X, additionally use some intervention to block indirect effect in some people (keep M from changing, set M to fixed level...)
    - if there is a (positive) indirect effect, then the effect of X on Y should be larger if the indirect effect isn't blocked
- this will obviously involve assumptions – for example, that we can block the indirect effect without changing the nature of the manipulation of X
- Ge (2023): Experimentally manipulating mediating processes

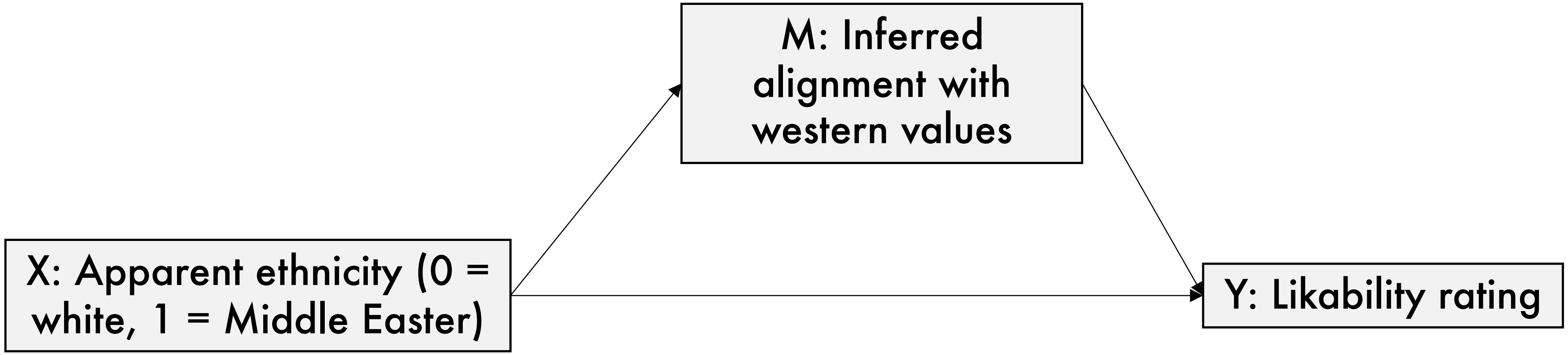
Simultaneous interventions

- In my (limited) review experience, researchers get confused when they try to map the mediation logic onto experimental interventions *in the abstract*
- I’d thus recommend to *initially* reason about this in terms of experimental design and hypotheses rather than in terms of mediation analysis
  - “If X affects Y via M, then if we do the following, we expect the following outcome…”
  - Makes better use of existing intuitions about threats to validity (which point towards necessary assumptions)
  - can still map onto mediation logic in the next step to more formally check whether you got everything right


M: Inferred alignment with western values
X: Apparent ethnicity (0 = white, 1 = Middle Easter)
Y: Likability rating

2 x 3 design:

- image of white or Middle Eastern person (manipulates X)
- additional information meant to manipulate M
  - none (M = “natural” value for X)
  - target volunteers for Amnesty international (M = high)
  - target volunteers for the Salafists (M = low)


M: Inferred alignment with western values
X: Apparent ethnicity (0 = white, 1 = Middle Easter)
Y: Likability rating

	"Natural"	Amnesty International	Salafists
White	$Y^{X = 0,\ M(X = 0)}$	$Y^{X = 0,\ M = \text{high}}$	$Y^{X = 0,\ M = \text{low}}$
Middle Eastern	$Y^{X = 1,\ M(X = 1)}$	$Y^{X = 1,\ M = \text{high}}$	$Y^{X = 1,\ M = \text{low}}$

Comparison yields total effect

Comparison yields *controlled* direct effect

Comparison yields a different *controlled* direct effect

We are *not* estimating natural effects here because we force certain values upon the mediator
→ we are *not* in the business of decomposing the total effect into additive parts

	“Natural”	Amnesty International	Salafists
White	$Y^{X = 0, M(X = 0)}$	$Y^{X = 0, M = \text{high}}$	$Y^{X = 0, M = \text{low}}$
Middle Eastern	$Y^{X = 1, M(X = 1)}$	$Y^{X = 1, M = \text{high}}$	$Y^{X = 1, M = \text{low}}$

Total effect ≈ Controlled direct effect → speaks against mediation
Total effect > Controlled direct effect → speaks in favor of mediation*
Total effect < Controlled direct effect → Could actually happen, (e.g., M*X interaction)

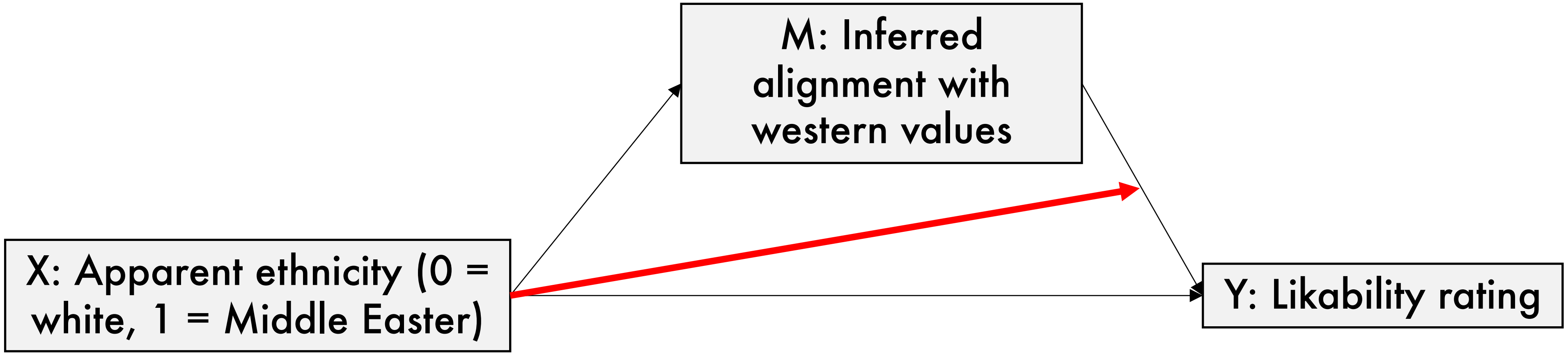
More precise conclusions will hinge on additional assumptions

Controlled direct effect ≈ 0 → full mediation?

“Threat to validity” of this conclusion:

- Volunteering for AI may mean different things depending on X
  - X = 0: not that extraordinary, X = 1: may speak to a special commitment/motivation
- M = high may thus give X = 1 a huge likability boost, which hides an existing direct effect
- at least 3 different ways to conceptualize this from a causal inference perspective

*see Ge (2023) for additional conditions (X affects measured M in the “natural” group, manipulation of M successfully changes measured M)

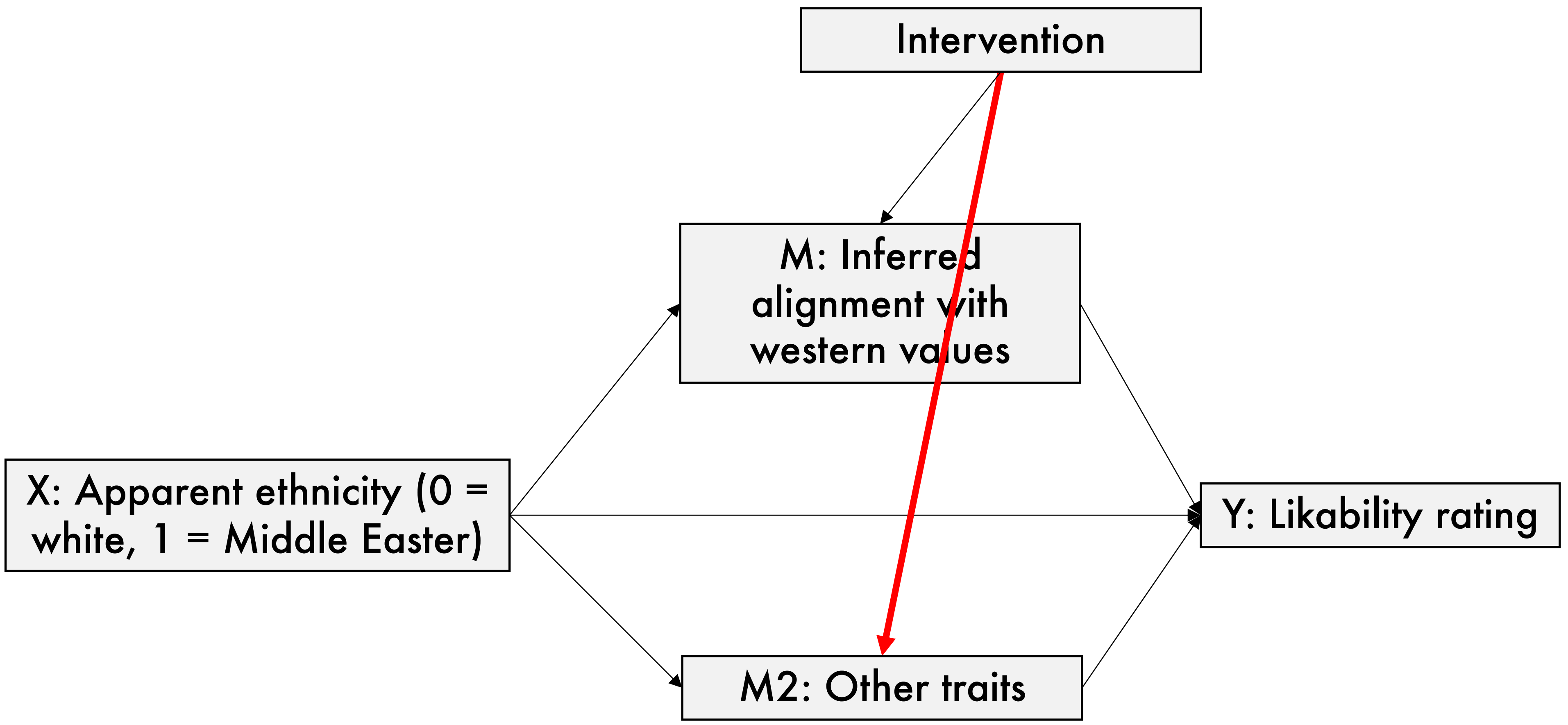


M: Inferred alignment with western values

X: Apparent ethnicity (0 = white, 1 = Middle Easter)

Y: Likability rating

Treatment-mediator interaction:
Effect of inferred values depends on apparent ethnicity


Intervention

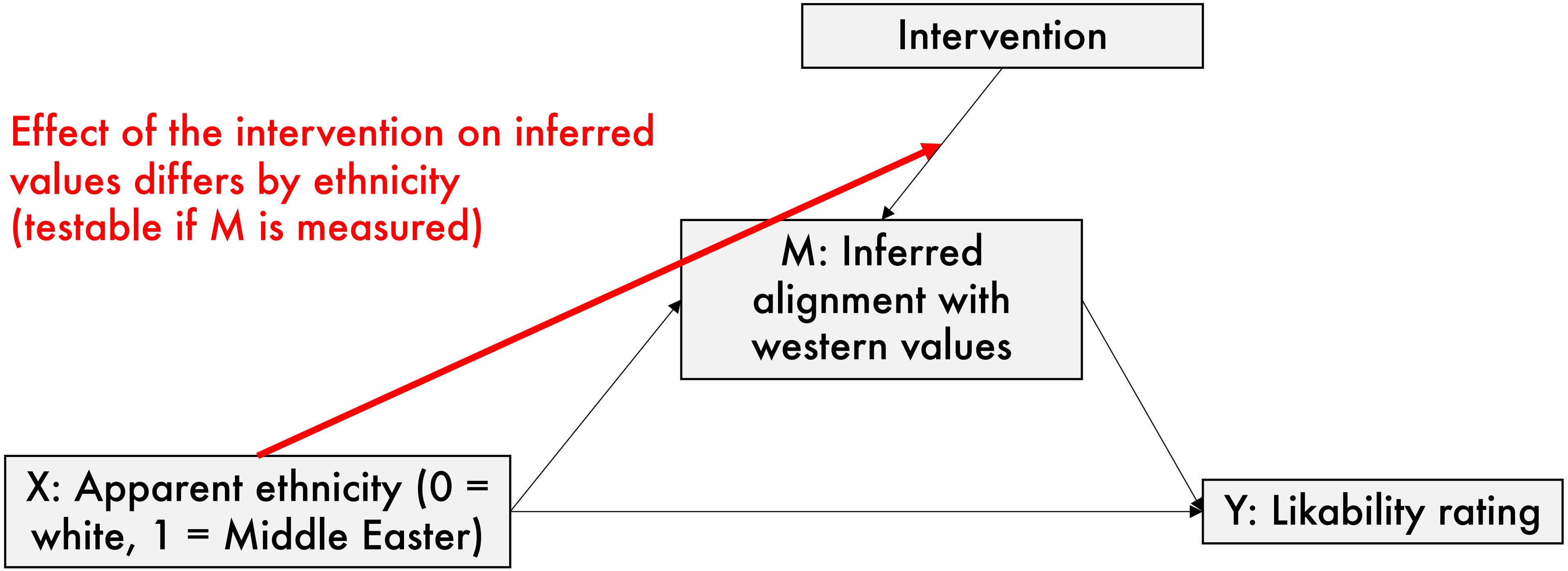
M: Inferred alignment with western values

X: Apparent ethnicity (0 = white, 1 = Middle Easter)

Y: Likability rating

M2: Other traits

Intervention affects another (possibly unobserved) mediator



Effect of the intervention on inferred values differs by ethnicity (testable if M is measured)

Intervention

M: Inferred alignment with western values

X: Apparent ethnicity (0 = white, 1 = Middle Easter)

Y: Likability rating

	"Natural"	Amnesty International	Salafists
White	$Y^{X = 0, M(X = 0)}$	$Y^{X = 0, M = \text{high}}$	$Y^{X = 0, M = \text{low}}$
Middle Eastern	$Y^{X = 1, M(X = 1)}$	$Y^{X = 1, M = \text{high}}$	$Y^{X = 1, M = \text{low}}$

Total effect ≈ Controlled direct effect → speaks against mediation
Total effect > Controlled direct effect → speaks in favor of mediation
Total effect < Controlled direct effect → Could actually happen, (e.g., M*X interaction)

More precise conclusions will hinge on additional assumptions

Controlled direct effect ≈ 0 → full mediation?

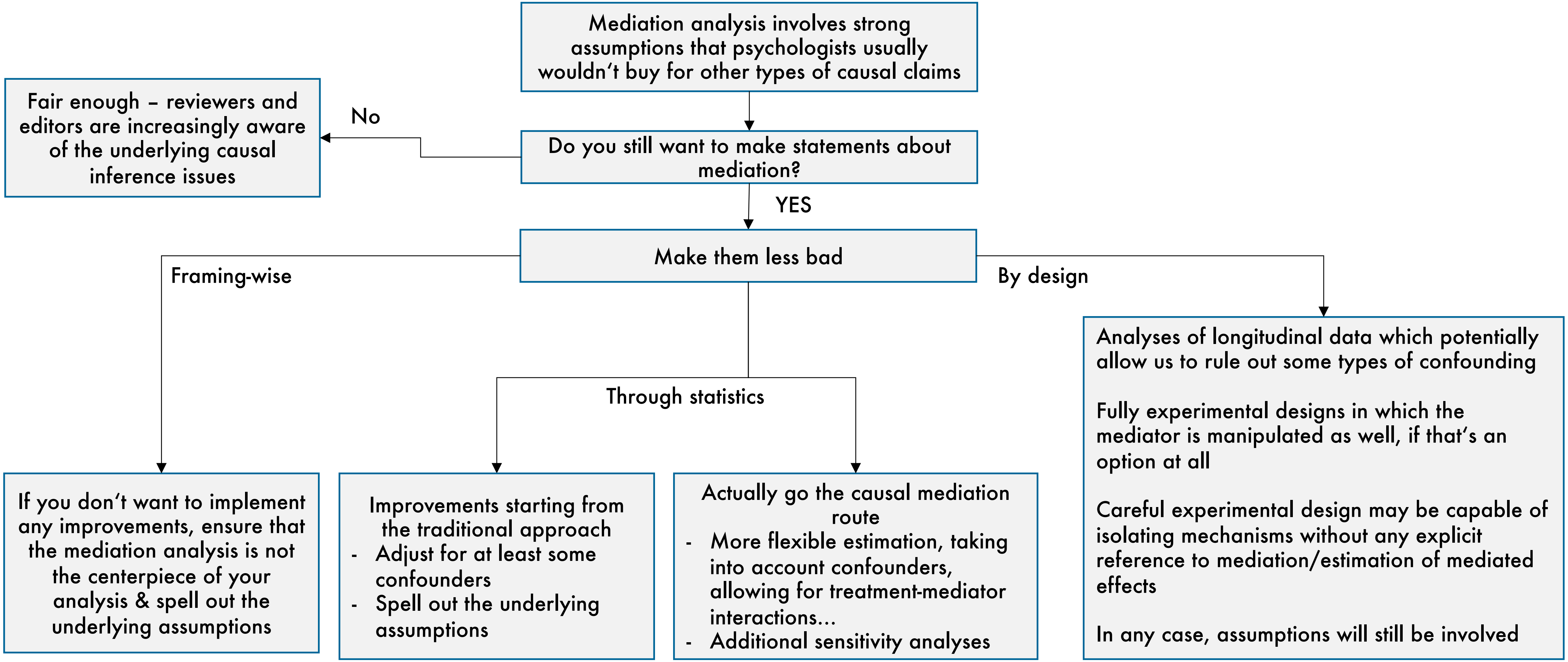
"Threat to validity" of this conclusion:

- Volunteering for AI may mean different things depending on X
  - X = 0: not that extraordinary, X = 1: may speak to a special commitment/motivation
- M = high may thus give X = 1 a huge likability boost, which hides a remaining direct effect
- different ways to conceptualize this from a causal inference perspective

These threats to validity must be assumed away to warrant the conclusion of full mediation

Matthay & Glymour (2020): A Graphical Catalog of Threats to Validity
Gabriel et al. (2025): Elucidating some common biases in randomized controlled trials using directed acyclic graphs

Okay, so that was a lot



Mediation analysis involves strong assumptions that psychologists usually wouldn't buy for other types of causal claims

Do you still want to make statements about mediation?

No

Fair enough – reviewers and editors are increasingly aware of the underlying causal inference issues

YES

Make them less bad

Framing-wise

If you don't want to implement any improvements, ensure that the mediation analysis is not the centerpiece of your analysis & spell out the underlying assumptions

Through statistics

Improvements starting from the traditional approach

- Adjust for at least some confounders
- Spell out the underlying assumptions

Actually go the causal mediation route

- More flexible estimation, taking into account confounders, allowing for treatment-mediator interactions…
- Additional sensitivity analyses

By design

Analyses of longitudinal data which potentially allow us to rule out some types of confounding

Fully experimental designs in which the mediator is manipulated as well, if that's an option at all

Careful experimental design may be capable of isolating mechanisms without any explicit reference to mediation/estimation of mediated effects

In any case, assumptions will still be involved

Thank you for your attention!

Julia M. Rohrer
Leipzig University
juliarohrer.com
www.the100.ci

Slides: Resources