Using Experiments to Test Process Theories

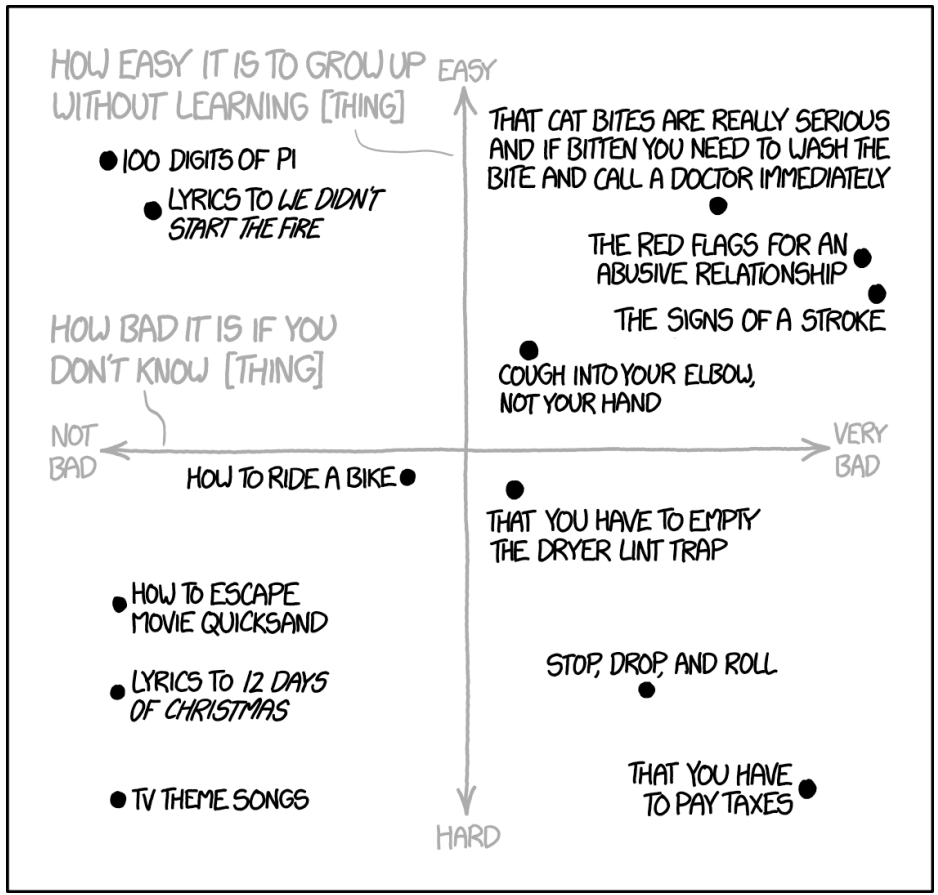
How and Why?

A talk based on Asay, Guggenmos, Kadous, Koonce, and Libby (forthcoming, TAR)

Ryan Guggenmos, Assistant Professor of Accounting

Johnson
Cornell
SC Johnson College of Business

Where are we at (learning-wise)?



(xkcd.com)

Today, I hope to..

1. Communicate the value of experimental research in business and economic contexts.

Today, I hope to..

1. Communicate the value of experimental research in business and economic contexts.

2. Enhance your ability to recognize the strengths and limitations of experimental research.

Today, I hope to..

1. Communicate the value of experimental research in business and economic contexts.

2. Enhance your ability to recognize the strengths and limitations of experimental research.

3. Develop your understanding of what makes a more or less effective experiment.

Today, I hope to..

1. Communicate the value of experimental research in business and economic contexts.

2. Enhance your ability to recognize the strengths and limitations of experimental research.

3. Develop your understanding of what makes a more or less effective experiment.

4. Show how experiments can help us better understand how and why causes become effects.

Disclaimer!

Experiments are not the only methodological approach that provides useful evidence.

Disclaimer!

Experiments are not the only methodological approach that provides useful evidence.

So why am I focusing on experiments…

Disclaimer!

Experiments are not the only methodological approach that provides useful evidence.

So why am I focusing on experiments…

- Well-designed experiments can provide strong causal evidence, that other approaches cannot.

Disclaimer!

Experiments are not the only methodological approach that provides useful evidence.

So why am I focusing on experiments…

- Well-designed experiments can provide strong causal evidence, that other approaches cannot.
- Experiments are suitable for testing nuanced theory and leveraging strict control to rule out alternative explanations.

Disclaimer!

Experiments are not the only methodological approach that provides useful evidence.

So why am I focusing on experiments…

- Well-designed experiments can provide strong causal evidence, that other approaches cannot.
- Experiments are suitable for testing nuanced theory and leveraging strict control to rule out alternative explanations.
- I am an experimental researcher, so I know more about experiments than other methods. (And I like them. :))

Applied Research

Research methods are the tools we use to test theories in the settings / contexts we care about.

For us, usually, these are business and economics contexts.

Applied Research

Research methods are the tools we use to test theories in the settings / contexts we care about.

For us, usually, these are business and economics contexts.

The usefulness of a "tool" is conditional on the goals of the project.

Applied Research

Research methods are the tools we use to test theories in the settings / contexts we care about.

For us, usually, these are business and economics contexts.

The usefulness of a "tool" is conditional on the goals of the project.

Example: If our goal is to provide strong causal evidence for how or why misreporting leads to poor firm performance, then we need a tool that is designed to help us provide this kind of evidence.

The Experimentalist's Advantage

The strength of a well-designed experiment lies in its ability to test a specific theory or set of theories related to a setting of interest.

Experiments

The Experimentalist's Advantage

The strength of a well-designed experiment lies in its ability to test a specific theory or set of theories related to a setting of interest.

Experiments

Strengths

- Testing for directional effects

The Experimentalist's Advantage

The strength of a well-designed experiment lies in its ability to test a specific theory or set of theories related to a setting of interest.

Experiments

Strengths

- Testing for directional effects
- Testing theory

The Experimentalist's Advantage

The strength of a well-designed experiment lies in its ability to test a specific theory or set of theories related to a setting of interest.

Experiments

Strengths

- Testing for directional effects
- Testing theory
- Internal validity

The Experimentalist's Advantage

The strength of a well-designed experiment lies in its ability to test a specific theory or set of theories related to a setting of interest.

Experiments

Strengths

- Testing for directional effects
- Testing theory
- Internal validity
- Causal inference

The Experimentalist's Advantage

The strength of a well-designed experiment lies in its ability to test a specific theory or set of theories related to a setting of interest.

	Strengths	Weaknesses
Experiments	• Testing for directional effects • Testing theory • Internal validity • Causal inference	• Estimating absolute levels

The Experimentalist's Advantage

The strength of a well-designed experiment lies in its ability to test a specific theory or set of theories related to a setting of interest.

	Strengths	Weaknesses
Experiments	• Testing for directional effects • Testing theory • Internal validity • Causal inference	• Estimating absolute levels • Modeling the “real-world”

The Experimentalist's Advantage

The strength of a well-designed experiment lies in its ability to test a specific theory or set of theories related to a setting of interest.

	Strengths	Weaknesses
Experiments	• Testing for directional effects • Testing theory • Internal validity • Causal inference	• Estimating absolute levels • Modeling the “real-world” • External Validity

The Experimentalist's Advantage

The strength of a well-designed experiment lies in its ability to test a specific theory or set of theories related to a setting of interest.

	Strengths	Weaknesses
Experiments	• Testing for directional effects • Testing theory • Internal validity • Causal inference	• Estimating absolute levels • Modeling the “real-world” • External Validity (not really!)

The Experimentalist's Advantage

The strength of a well-designed experiment lies in its ability to test a specific theory or set of theories related to a setting of interest.

	Strengths	Weaknesses
Experiments	• Testing for directional effects • Testing theory • Internal validity • Causal inference	• Estimating absolute levels • Modeling the “real-world” • External Validity (not really!)

Theories vary on many dimensions, but experiments are especially good at providing theory-consistent *causal* evidence and testing "how", "when", or "why" causes become effects.

When is evidence causal?

1. The cause preceeded the effect.

2. The cause was related to the effect.

3. No plausible alternative explanations exist for the effect.

Shadish, Cook, and Campbell 2002

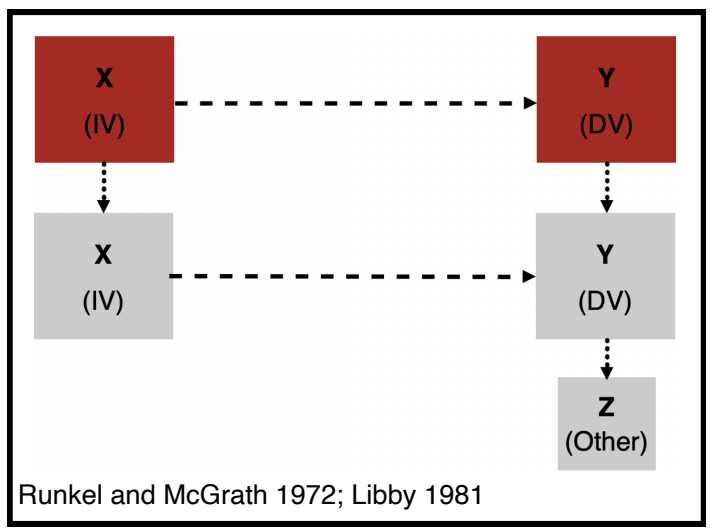
What makes a good design?

The goals:

1. Create a high-quality, but imperfect source of counterfactual variation.

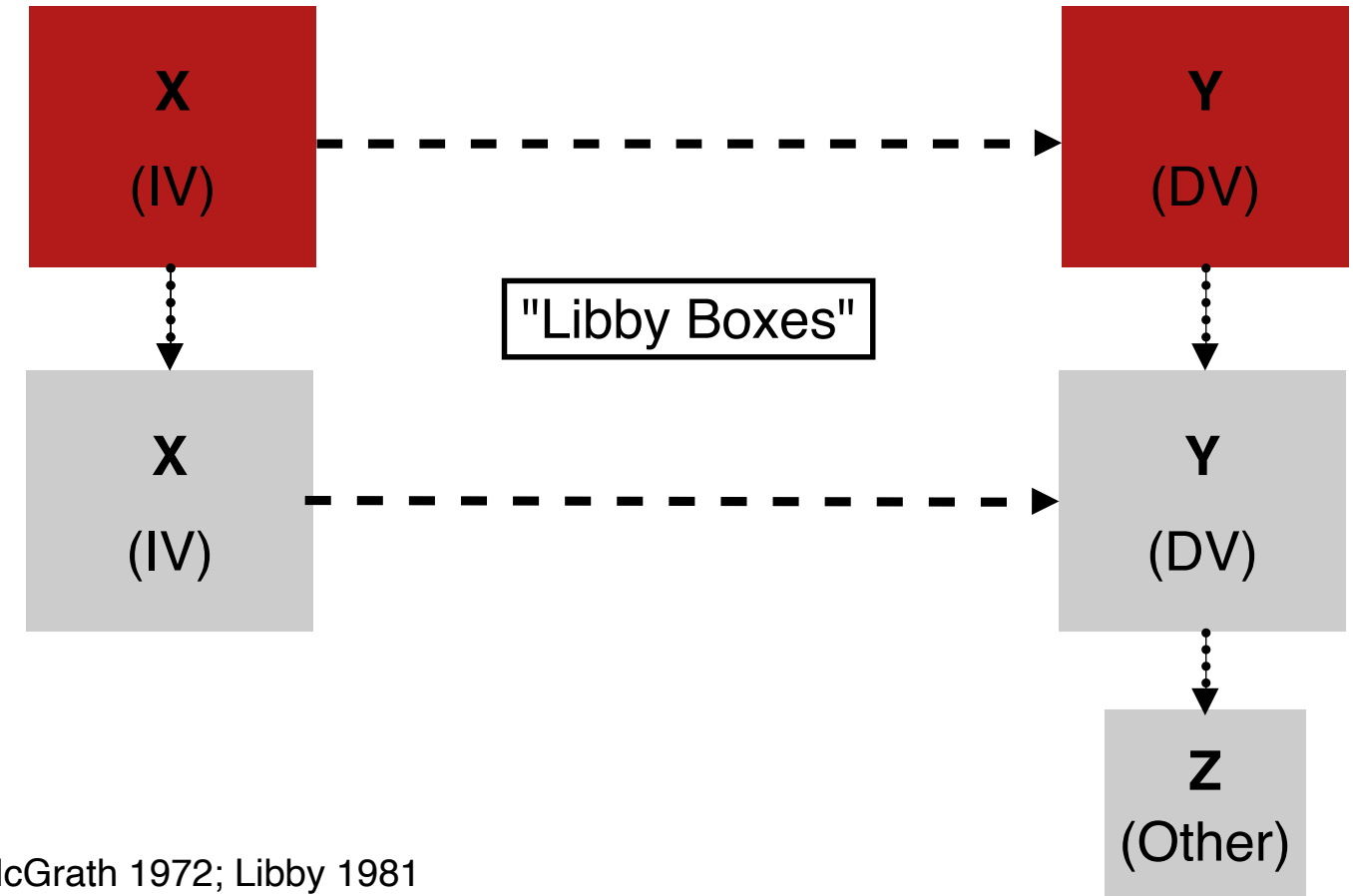
2. Understand how this source differs from the treatment condition.

Shadish, Cook, and Campbell 2002



Runkel and McGrath 1972; Libby 1981

The Predictive Validity Framework



Runkel and McGrath 1972; Libby 1981

When is evidence causal?

Shadish, Cook, and Campbell 2002

1. The cause preceeded the effect.

2. The cause was related to the effect.

3. No plausible alternative explanations for the effect.

Shadish, Cook, and Campbell 2002

When is evidence causal?

1. The cause preceeded the effect.

2. The cause was related to the effect.

3. No plausible alternative explanations for the effect.

How do experiments do this?

Shadish, Cook, and Campbell 2002

When is evidence causal?

1. The cause preceeded the effect.

2. The cause was related to the effect.

3. No plausible alternative explanations for the effect.

How do experiments do this?

1. Manipulate the cause and observe an outcome afterward.

2. See whether variation in cause is related to variation in effect.

3. Use design to rule out alternative explanations (usually through randomization and holding other variables constant).

Process Theories and Evidence

What are process theories?

> Process theories are stories related to
>
> *how*, *when, and why*
>
> causes become effects.

Process Theories and Evidence

What are process theories?

> Process theories are stories related to
>
> *how*, *when, and why*
>
> causes become effects.

What is process evidence?

Process Theories and Evidence

What are process theories?

> Process theories are stories related to
>
> *how*, *when, and why*
>
> causes become effects.

What is process evidence?

> Process evidence helps to refute, support, qualify or further our understanding of process theories.
>
> Process evidence can come in many forms.

Process Theories and Evidence

What are process theories?

> Process theories are stories related to
>
> *how*, *when, and why*
>
> causes become effects.

What is process evidence?

> Process evidence helps to refute, support, qualify or further our understanding of process theories.
>
> Process evidence can come in many forms.

Should we care about process? If so, when?

Process Theories and Evidence

What are process theories?

Process theories are stories related to

how, *when, and why*

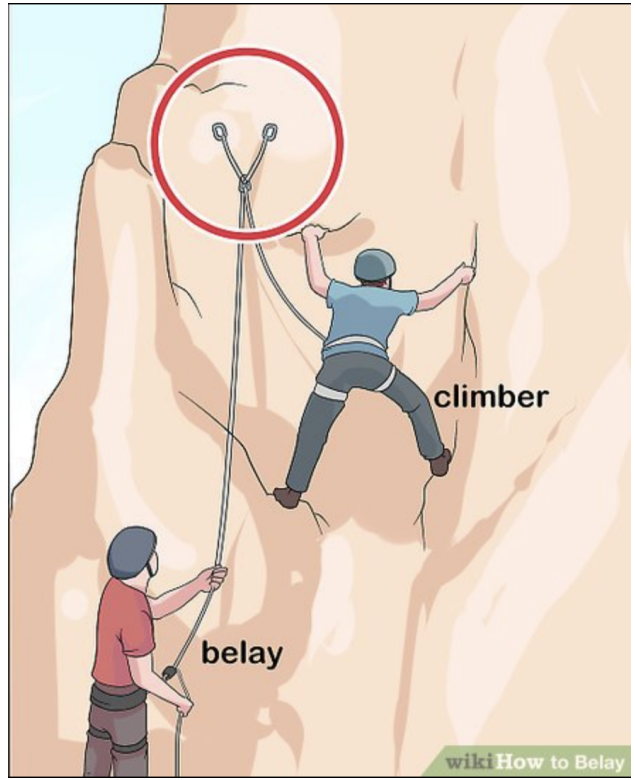
causes become effects.

What is process evidence?

Process evidence helps to refute, support, qualify or further our understanding of process theories.

Process evidence can come in many forms.

Should we care about process? If so, when? Yes, when it matters.



Belay device (ATC)


wikiHow to Belay


wikiHow to Belay


wikiHow to Belay


wikiHow to Belay


wikiHow to Belay


wikiHow to Belay

How does it work?

How does it work?

- Rope Spikes?

How does it work?

- Rope Spikes?
- Friction?

How does it work?

- Rope Spikes?
- Friction?
- Magnets?

How does it work?

- Rope Spikes?
- Friction?
- Magnets?
- Divine Intervention?

Process, Experiments, and Validity

How are these three related?

	Strengths	Weaknesses
Experiments	• Testing for directional effects • Testing theory • Internal validity • Causal inference	• Estimating absolute levels • Modeling the “real-world” • External Validity (not really!)

Process, Experiments, and Validity

How are these three related?

	Strengths	Weaknesses
Experiments	• Testing for directional effects • Testing theory • Internal validity • Causal inference	• Estimating absolute levels • Modeling the “real-world” • External Validity (not really!)

But wait! Process theories are theories…

Process, Experiments, and Validity

How are these three related?

	Strengths	Weaknesses
Experiments	• Testing for directional effects • Testing theory • Internal validity • Causal inference	• Estimating absolute levels • Modeling the “real-world” • External Validity (not really!)

But wait! Process theories are theories…

...And theory is what generalizes from an experiment.

Process, Experiments, and Validity

How are these three related?

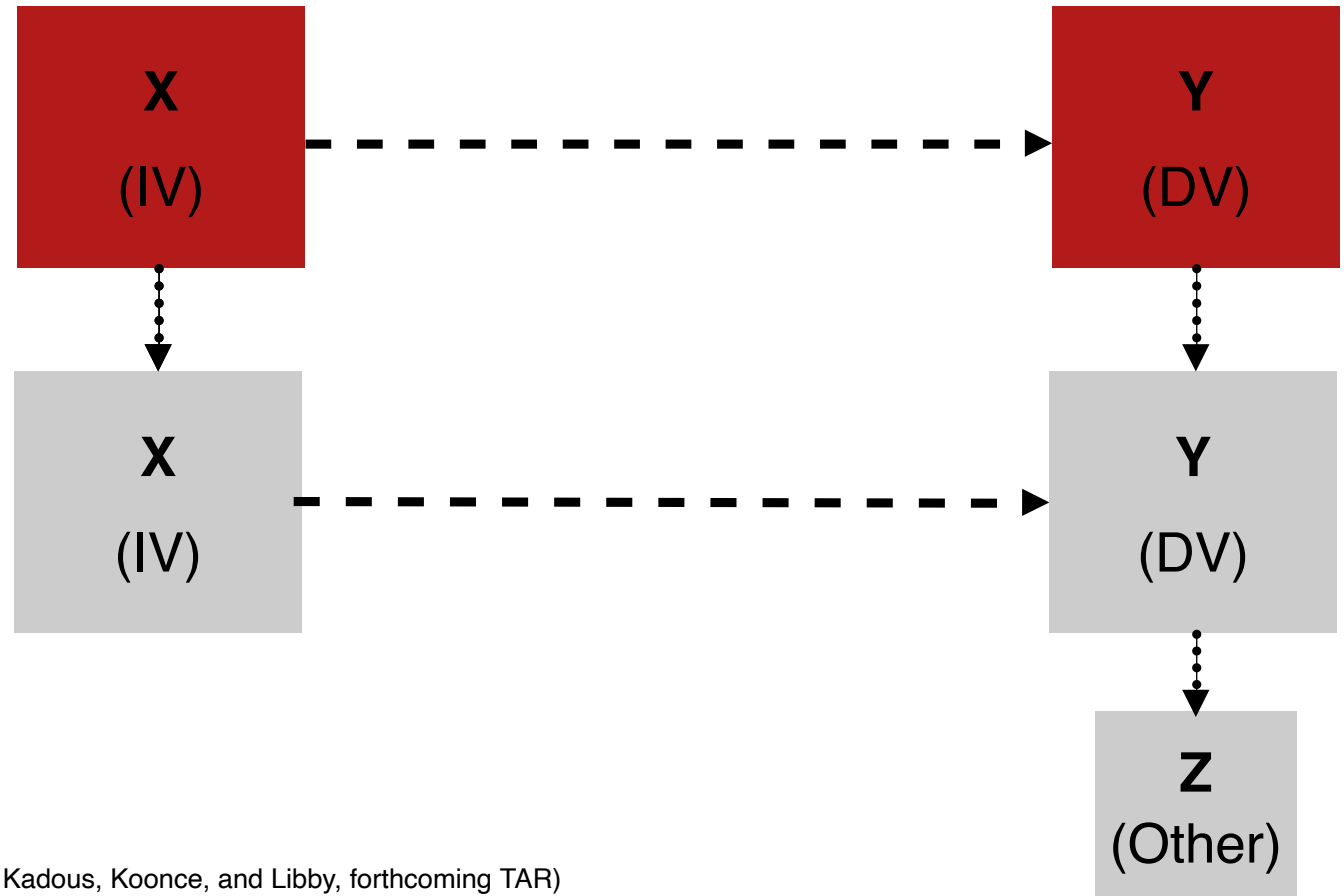
	Strengths	Weaknesses
Experiments	• Testing for directional effects • Testing theory • Internal validity • Causal inference	• Estimating absolute levels • Modeling the “real-world” • External Validity (not really!)

But wait! Process theories are theories…

...And theory is what generalizes from an experiment.

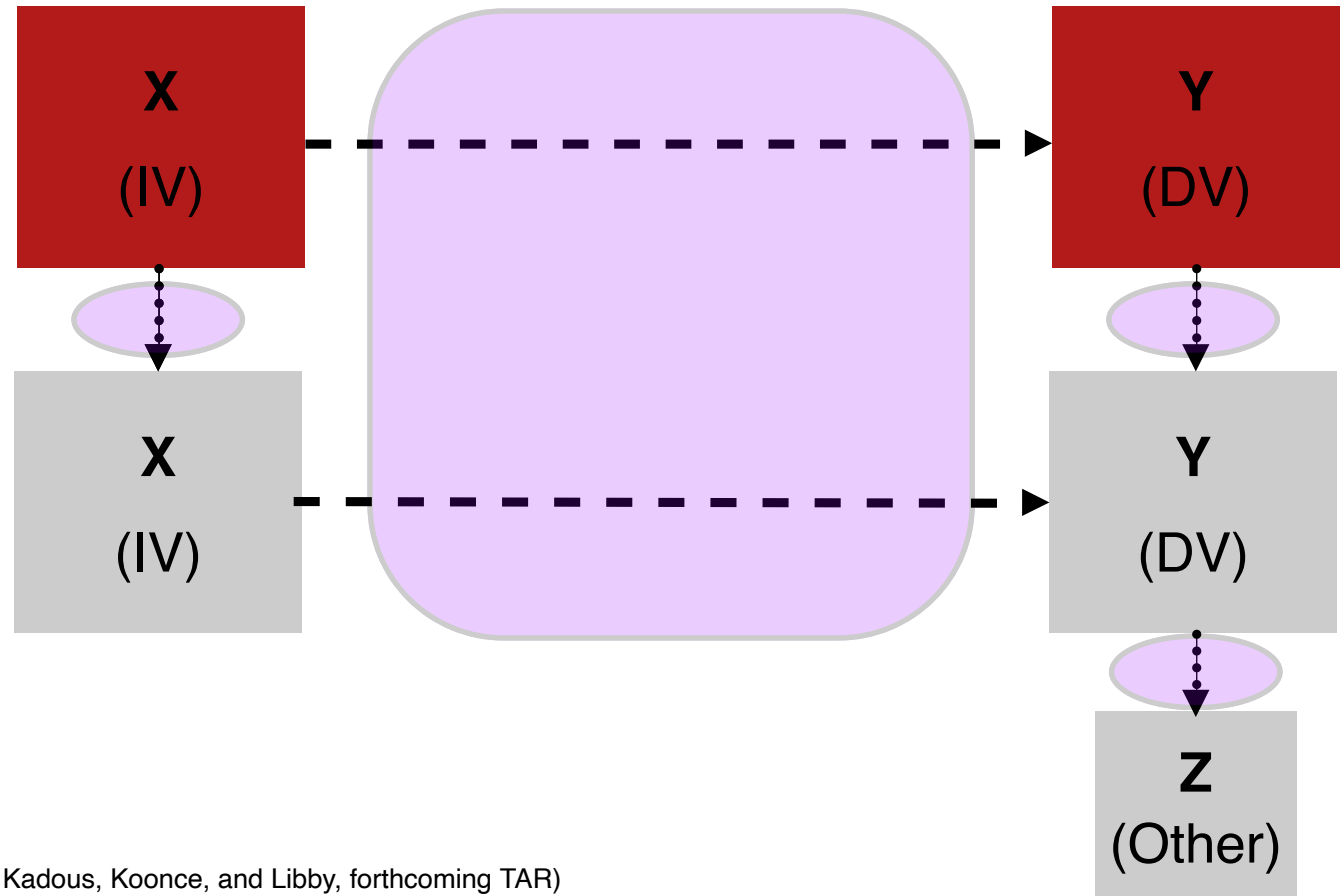
Convincing process evidence helps us generalize from our experiment to the "real world".

Let's get more complicated.



(Asay, Guggenmos, Kadous, Koonce, and Libby, forthcoming TAR)

Let's get more complicated.



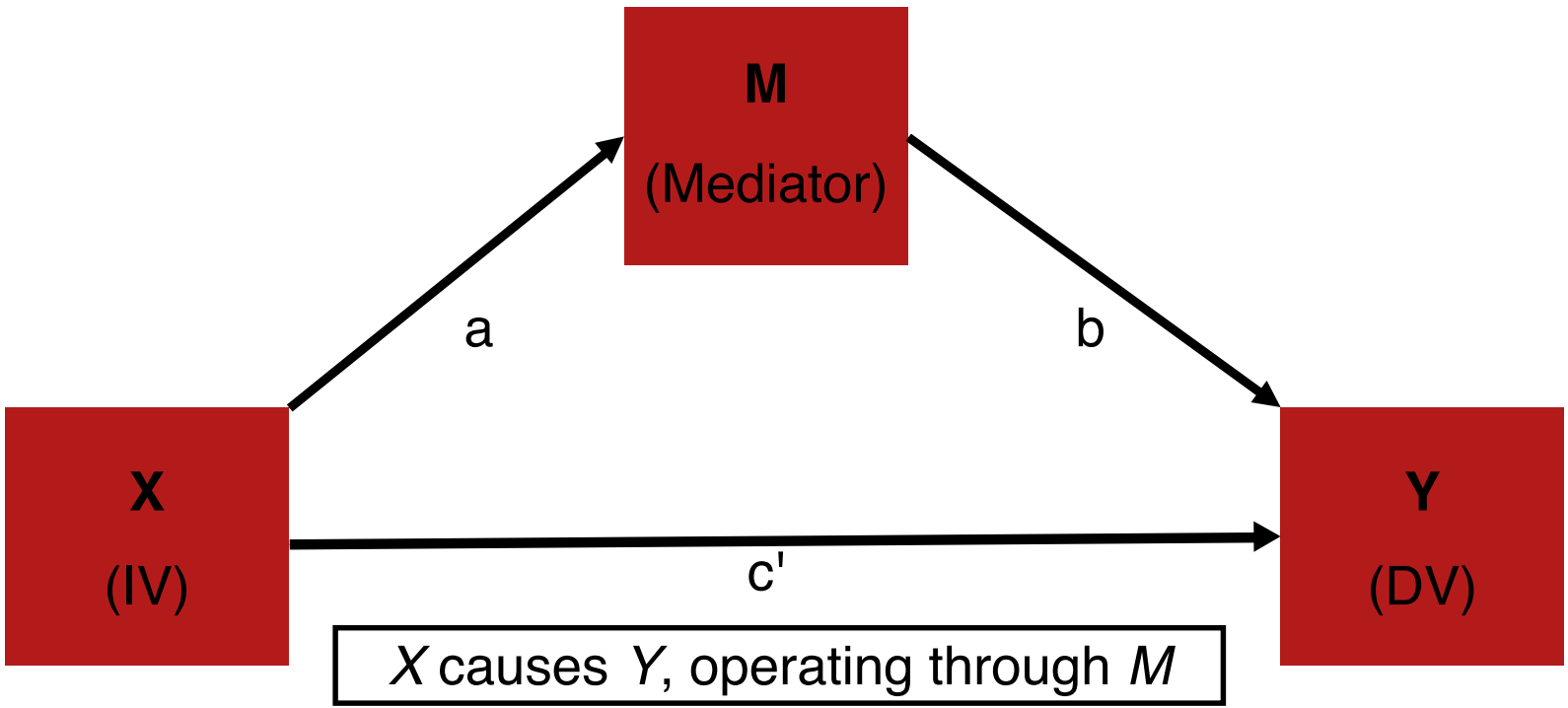
(Asay, Guggenmos, Kadous, Koonce, and Libby, forthcoming TAR)

Mediation

Mediation helps us understand *how or why* causes become effects.

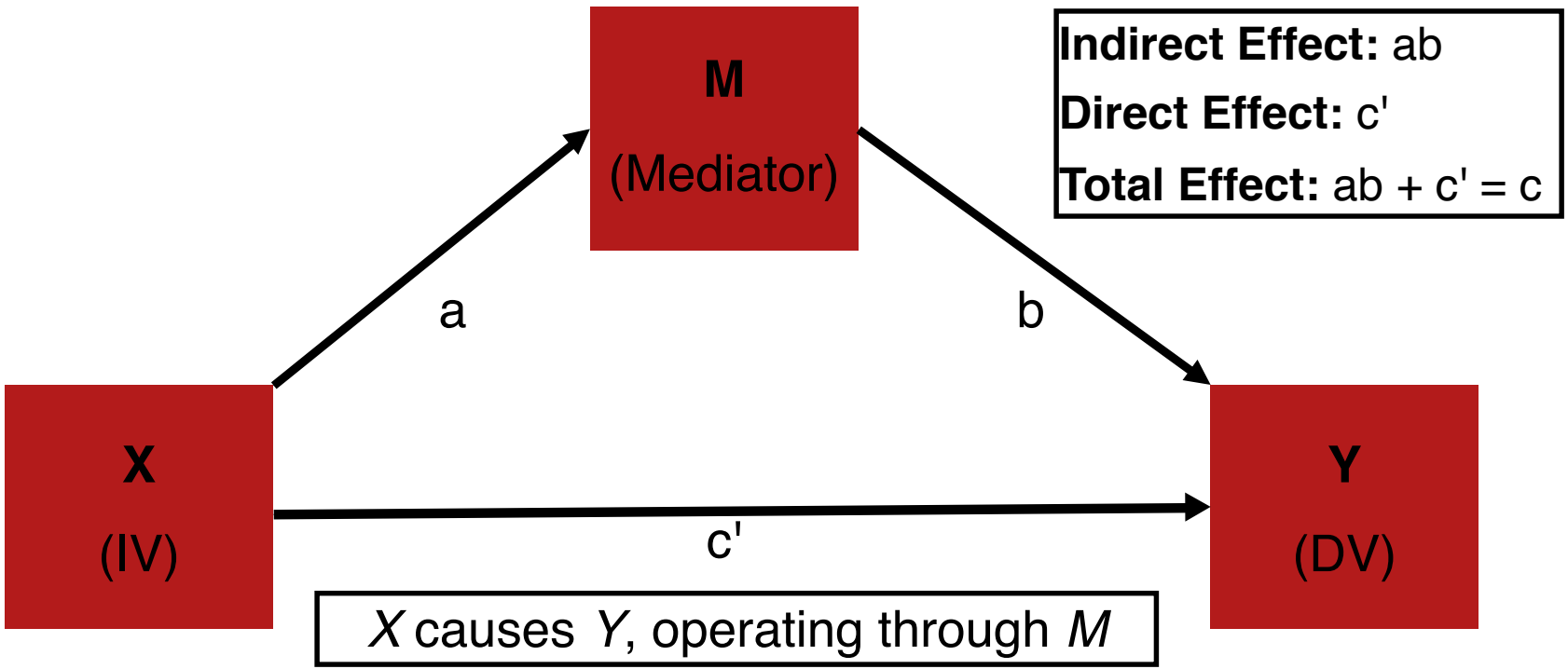
Mediation

Mediation helps us understand *how or why* causes become effects.



Mediation

Mediation helps us understand *how or why* causes become effects.



Mediation

Mediation can help us understand *how or why* causes become effects.

Mediation

Mediation can help us understand *how or why* causes become effects.

Strengths

Mediation

Mediation can help us understand *how or why* causes become effects.

Strengths

- Can provide convincing *correlational* process evidence.
- Generally allows for efficient designs.
- Usually easy to implement.
- Can be implemented without altering experimental flow.
- Analysis is (usually) straightforward with modern stats software.

Mediation

Mediation can help us understand *how or why* causes become effects.

Strengths

- Can provide convincing *correlational* process evidence.
- Generally allows for efficient designs.
- Usually easy to implement.
- Can be implemented without altering experimental flow.
- Analysis is (usually) straightforward with modern stats software.

Weaknesses

Mediation

Mediation can help us understand *how or why* causes become effects.

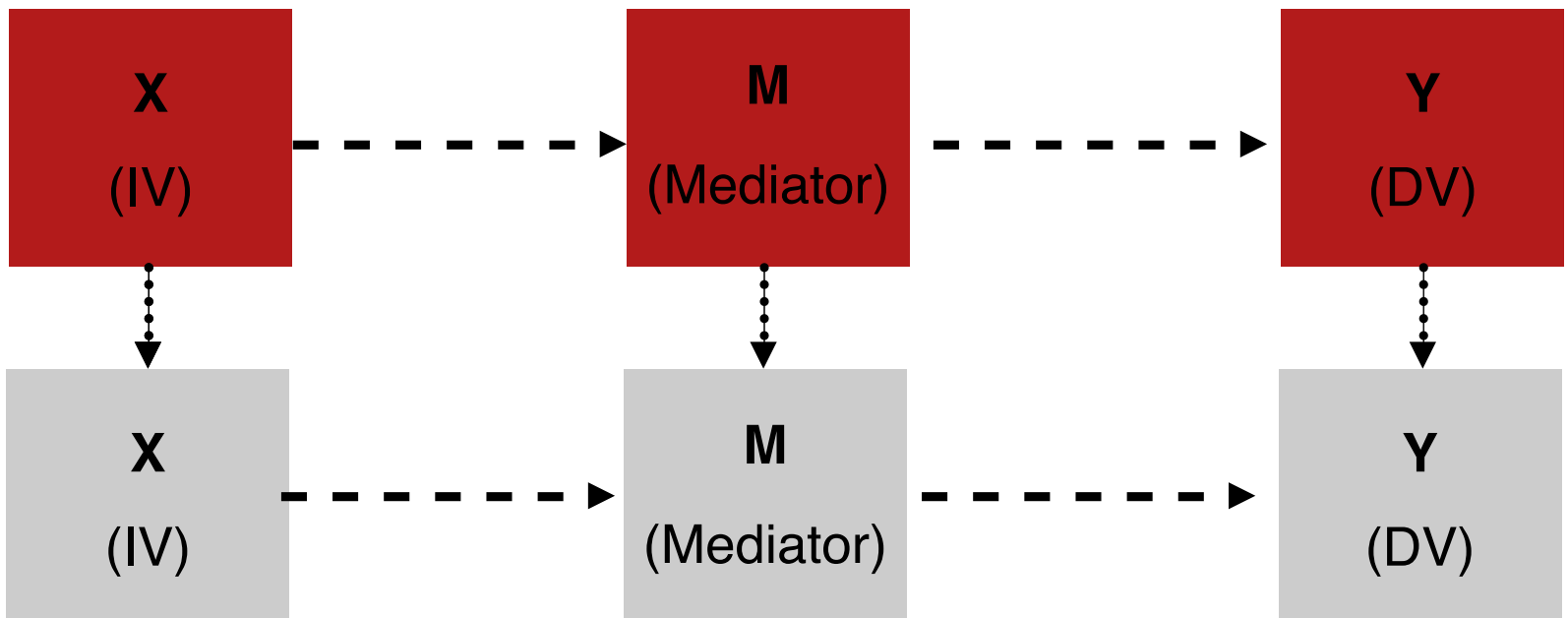
Strengths

- Can provide convincing *correlational* process evidence.
- Generally allows for efficient designs.
- Usually easy to implement.
- Can be implemented without altering experimental flow.
- Analysis is (usually) straightforward with modern stats software.

Weaknesses

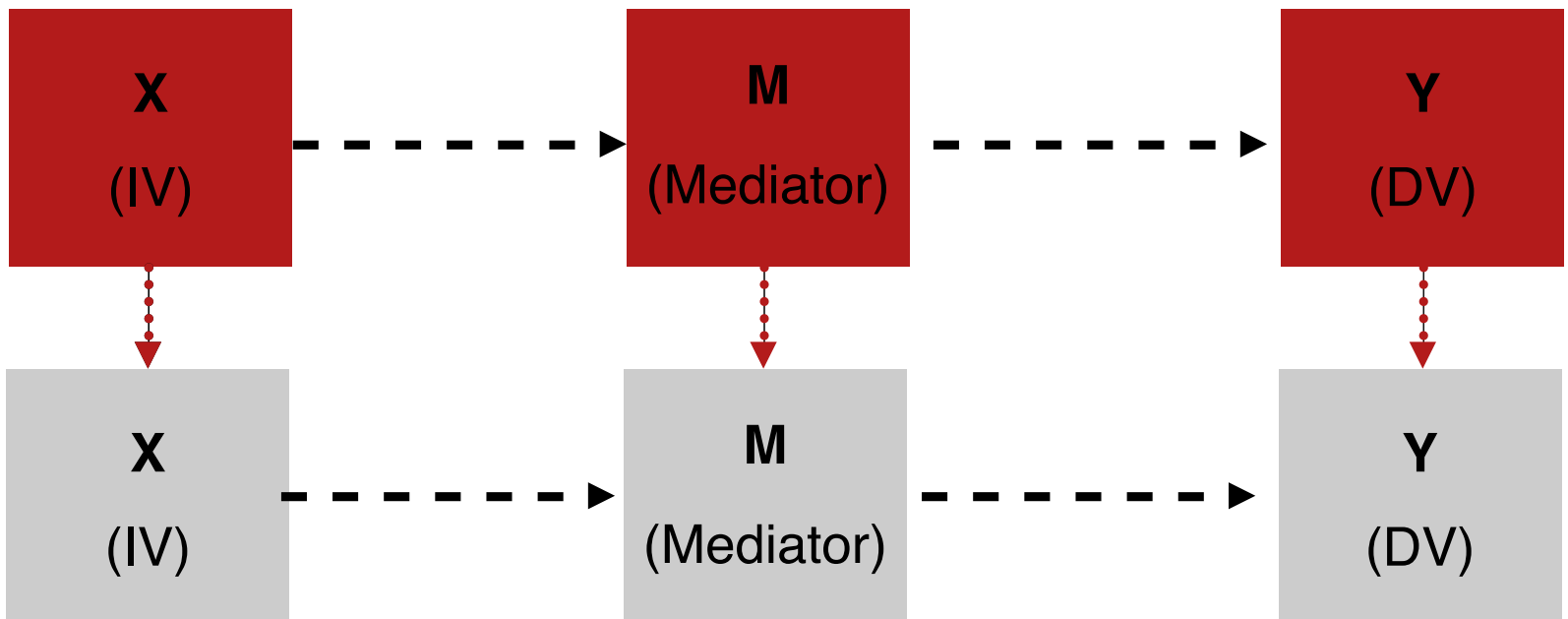
- Can only provide *correlational*, not causal evidence.
- Often relies on PEQs or participant self-insight, which can introduce validity threats.
- Relies on measured variables, so subject to measurement error concerns.
- No benefit of random assignment to the mediator.
- Coefficients are often biased.

Mediation: What could go wrong?



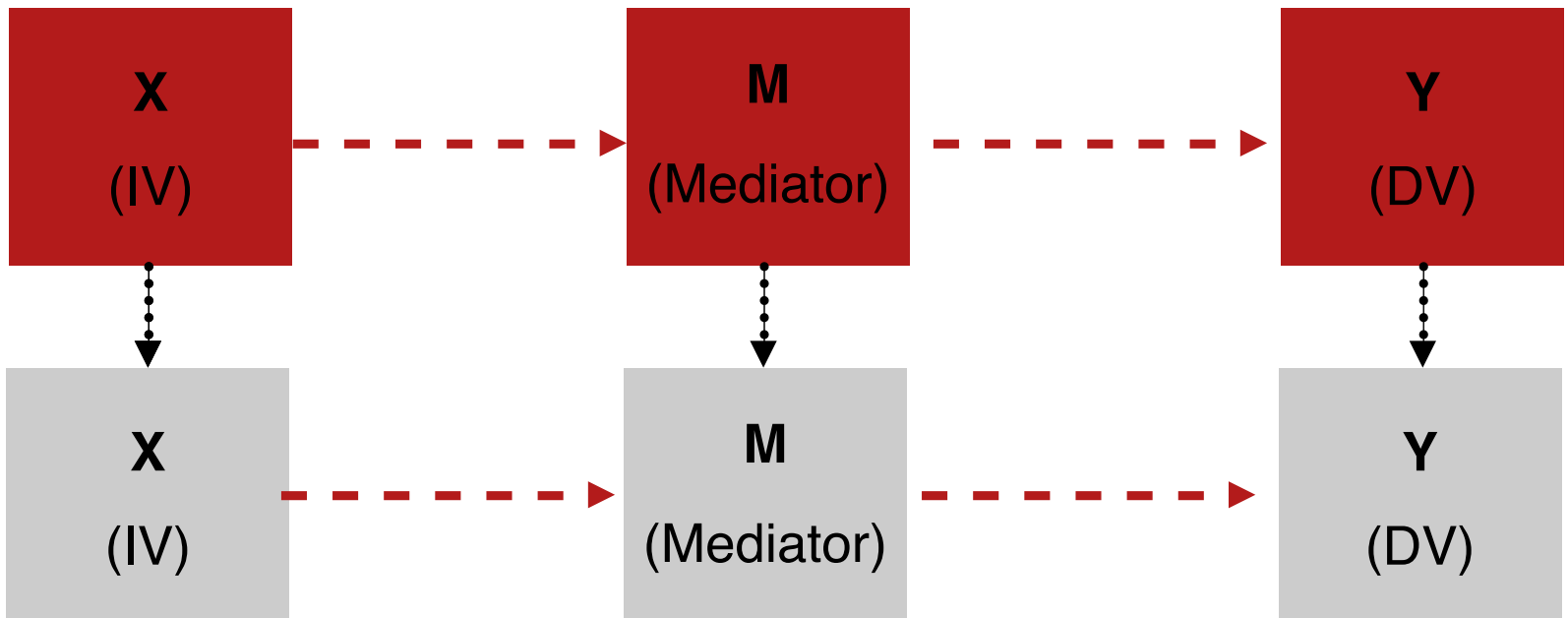
Mediation: What could go wrong?

Construct Operationalization Threats



Mediation: What could go wrong?

Variable Sequence Threats



Mediation: Best Practices

Mediation: Best Practices

How to Measure

Mediation: Best Practices

How to Measure

- Unobtrusive Measures
  - Technology
  - Coded Responses

Mediation: Best Practices

How to Measure

- Unobtrusive Measures
  - Technology
  - Coded Responses
- Obtrusive Measures
  - Scale and Question Design

Mediation: Best Practices

How to Measure

- Unobtrusive Measures
  - Technology
  - Coded Responses
- Obtrusive Measures
  - Scale and Question Design
- Participant Self-insight

Mediation: Best Practices

How to Measure

When to Measure

- Unobtrusive Measures
  - Technology
  - Coded Responses
- Obtrusive Measures
  - Scale and Question Design
- Participant Self-insight

Mediation: Best Practices

How to Measure

- Unobtrusive Measures
  - Technology
  - Coded Responses
- Obtrusive Measures
  - Scale and Question Design
- Participant Self-insight

When to Measure

- Carryover from *M* to *Y*
- Carryover from *Y* to *M*
- Carryover from *M* to other *M*s.

Mediation: Best Practices

How to Measure

- Unobtrusive Measures
  - Technology
  - Coded Responses
- Obtrusive Measures
  - Scale and Question Design
- Participant Self-insight

When to Measure

- Carryover from *M* to *Y*
- Carryover from *Y* to *M*
- Carryover from *M* to other *M*s.

What to Measure

Mediation: Best Practices

How to Measure

- Unobtrusive Measures
  - Technology
  - Coded Responses
- Obtrusive Measures
  - Scale and Question Design
- Participant Self-insight

When to Measure

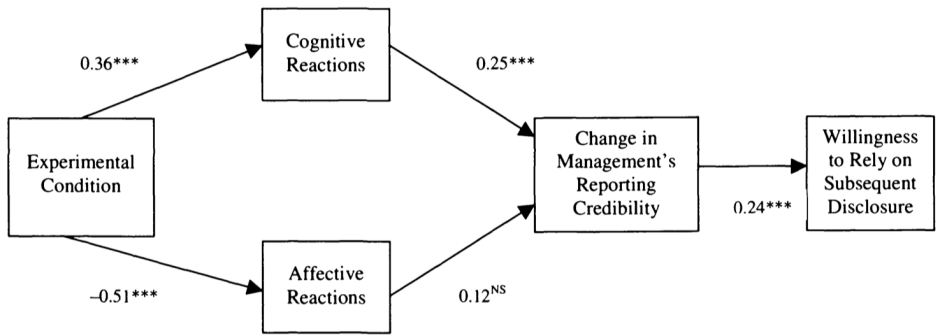
- Carryover from *M* to *Y*
- Carryover from *Y* to *M*
- Carryover from *M* to other *M*s.

What to Measure

- One "thing", all "things", or somewhere in-between?

Credibility and forthcomingness

Panel A: Determinants and Consequences of Reporting Credibility in the Short-Term



Panel B: Determinants and Consequences of Reporting Credibility in the Long-Term

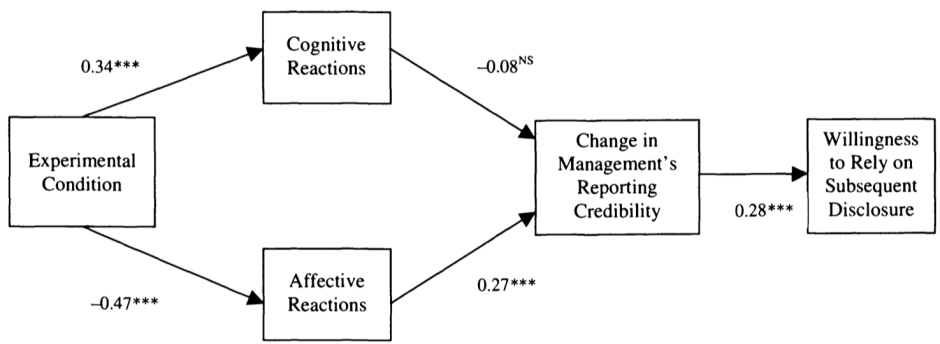
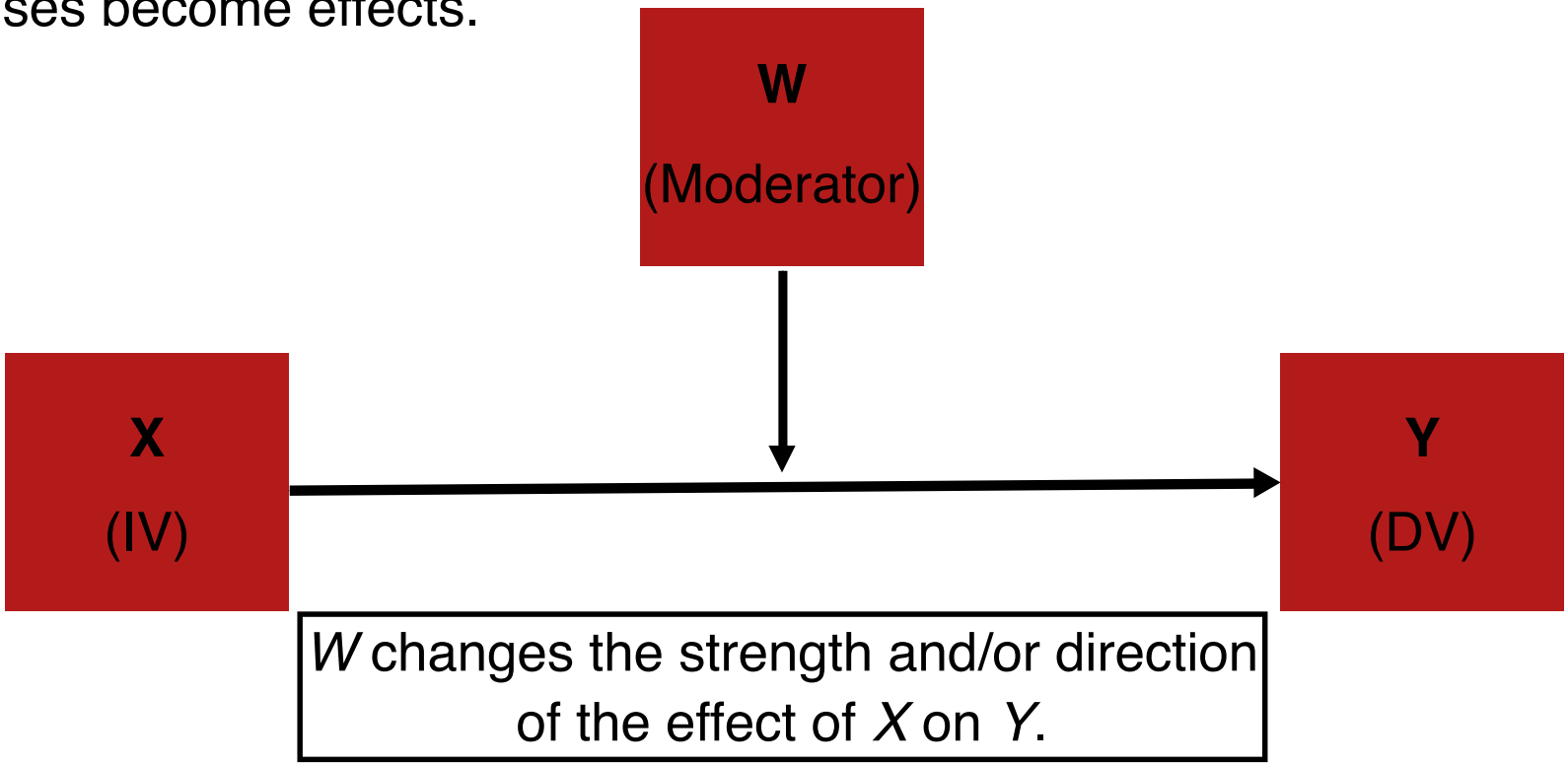


Figure 3 - Mercer (2005 TAR)

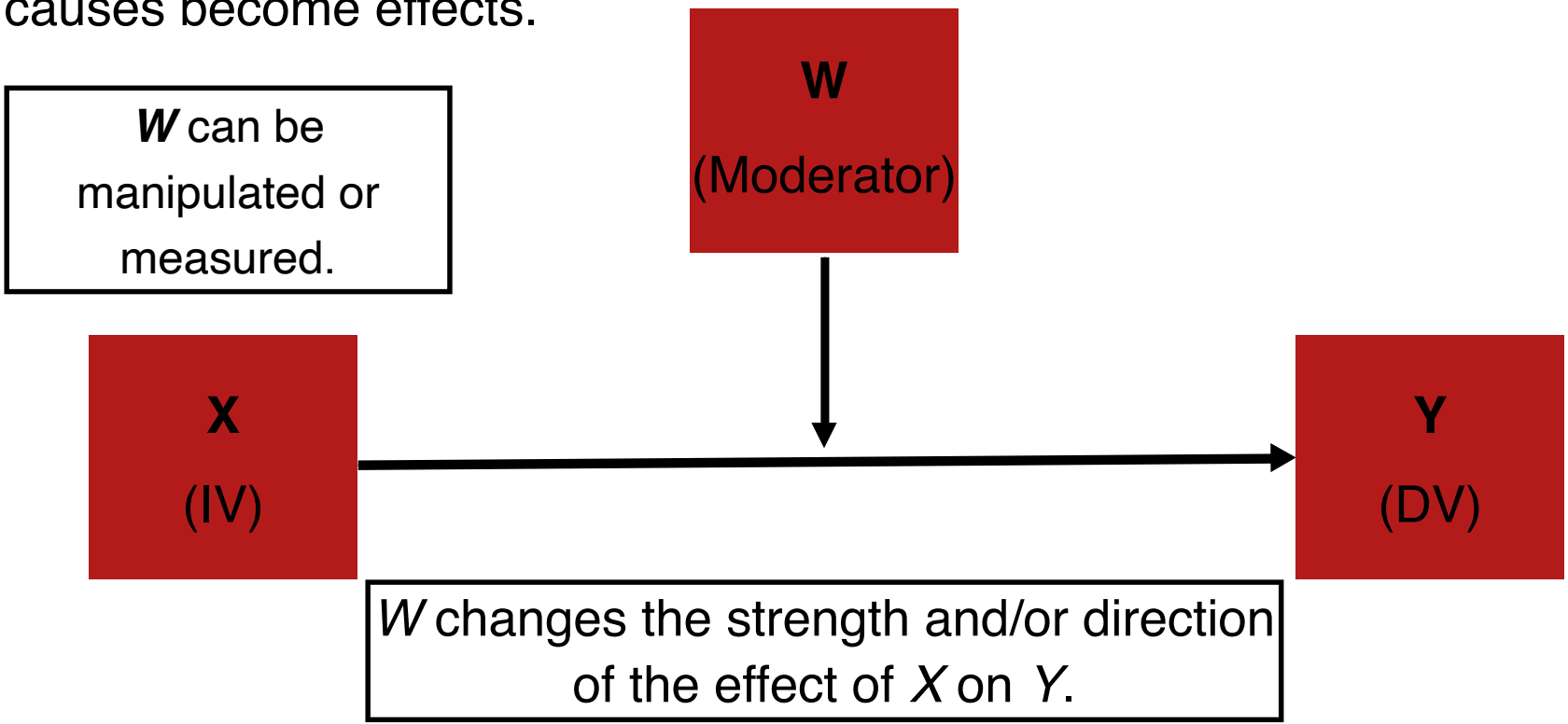
Moderation

Moderation helps us understand *when (or under what circumstances)* causes become effects.



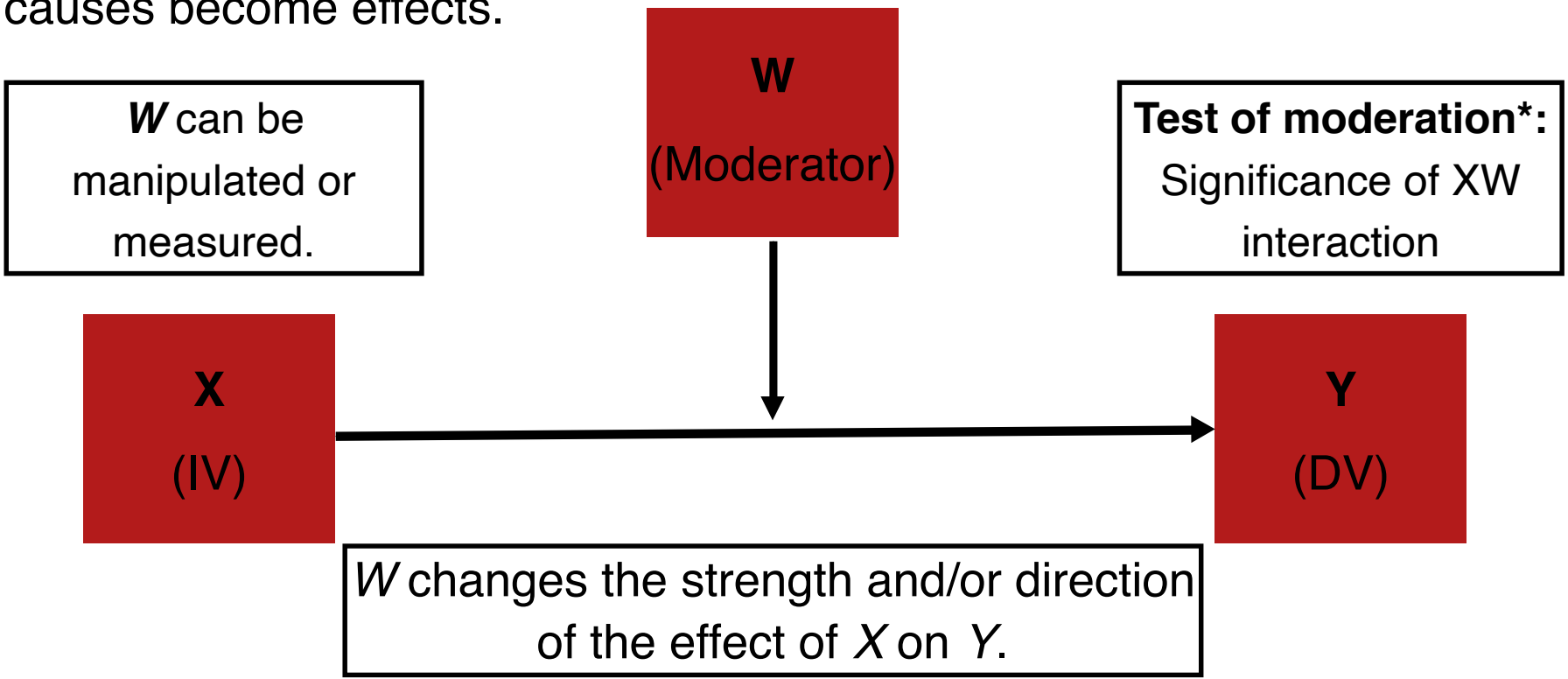
Moderation

Moderation helps us understand *when (or under what circumstances)* causes become effects.



Moderation

Moderation helps us understand *when (or under what circumstances)* causes become effects.



Moderation

Moderation can help us understand *when (or under what circumstances)* causes become effects.

Moderation

Moderation can help us understand *when (or under what circumstances)* causes become effects.

Strengths

Moderation

Moderation can help us understand *when (or under what circumstances)* causes become effects.

Strengths

- Can provide convincing *causal* process evidence (only when the moderator is manipulated).
- Manipulated moderators can leverage random assignment and control of temporal order to rule out plausible alternative theories.
- Does not (usually) rely on fancier stats than an ANOVA or regression.
- Manipulated moderators largely alleviate measurement error concerns.
- Often provide confirmatory and disconfirmatory evidence at the same time.

Moderation

Moderation can help us understand *when (or under what circumstances)* causes become effects.

Strengths

- Can provide convincing *causal* process evidence (only when the moderator is manipulated).
- Manipulated moderators can leverage random assignment and control of temporal order to rule out plausible alternative theories.
- Does not (usually) rely on fancier stats than an ANOVA or regression.
- Manipulated moderators largely alleviate measurement error concerns.
- Often provide confirmatory and disconfirmatory evidence at the same time.

Weaknesses

Moderation

Moderation can help us understand *when (or under what circumstances)* causes become effects.

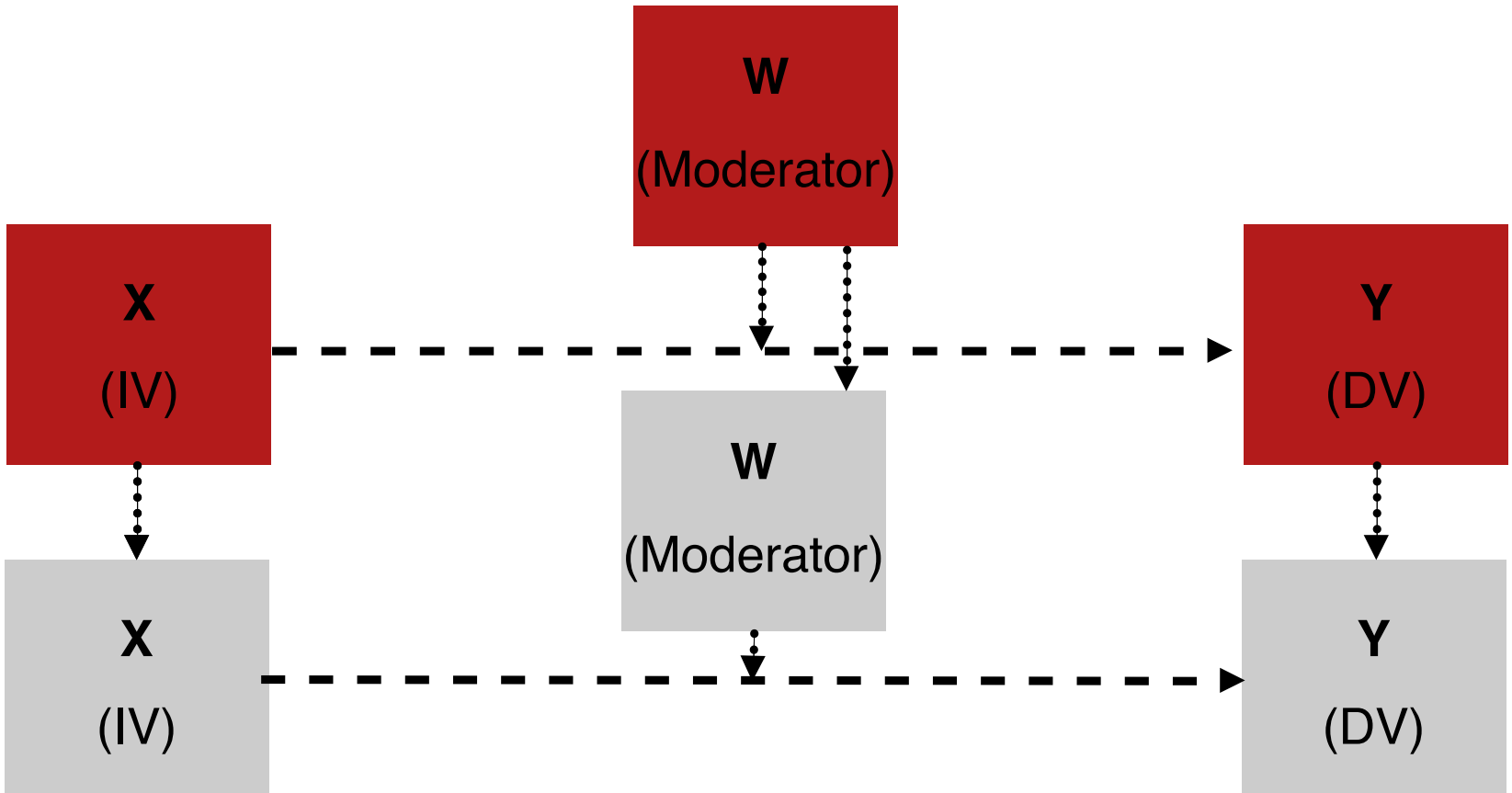
Strengths

- Can provide convincing *causal* process evidence (only when the moderator is manipulated).
- Manipulated moderators can leverage random assignment and control of temporal order to rule out plausible alternative theories.
- Does not (usually) rely on fancier stats than an ANOVA or regression.
- Manipulated moderators largely alleviate measurement error concerns.
- Often provide confirmatory and disconfirmatory evidence at the same time.

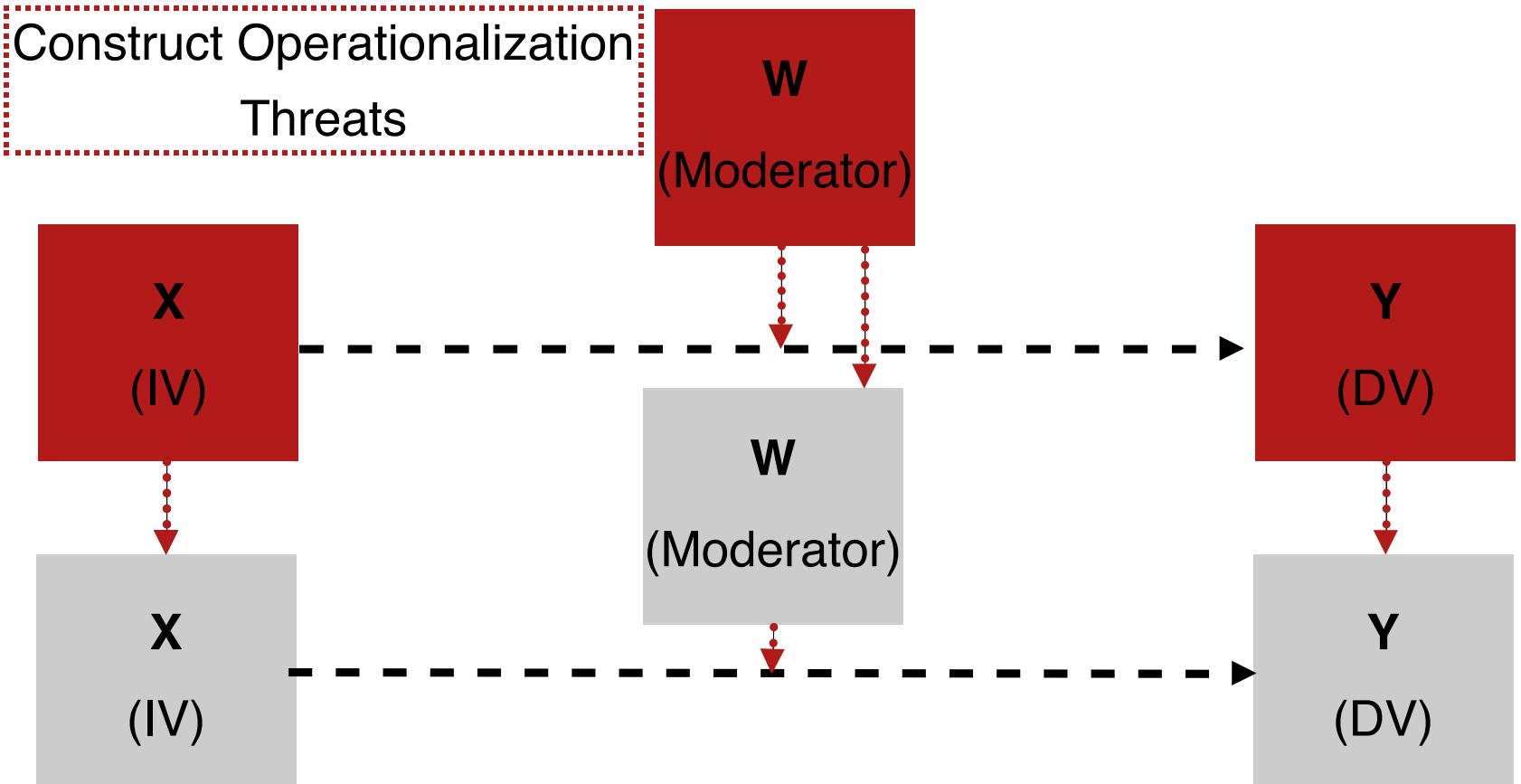
Weaknesses

- Many of the weaknesses of mediation also apply to measured moderator designs.
- Manipulated moderator designs often entail high participant resource demands (less efficient).
- Some moderators are not able to be (ethically) manipulated.
- Moderators are often more difficult to design (more abstracted from process).

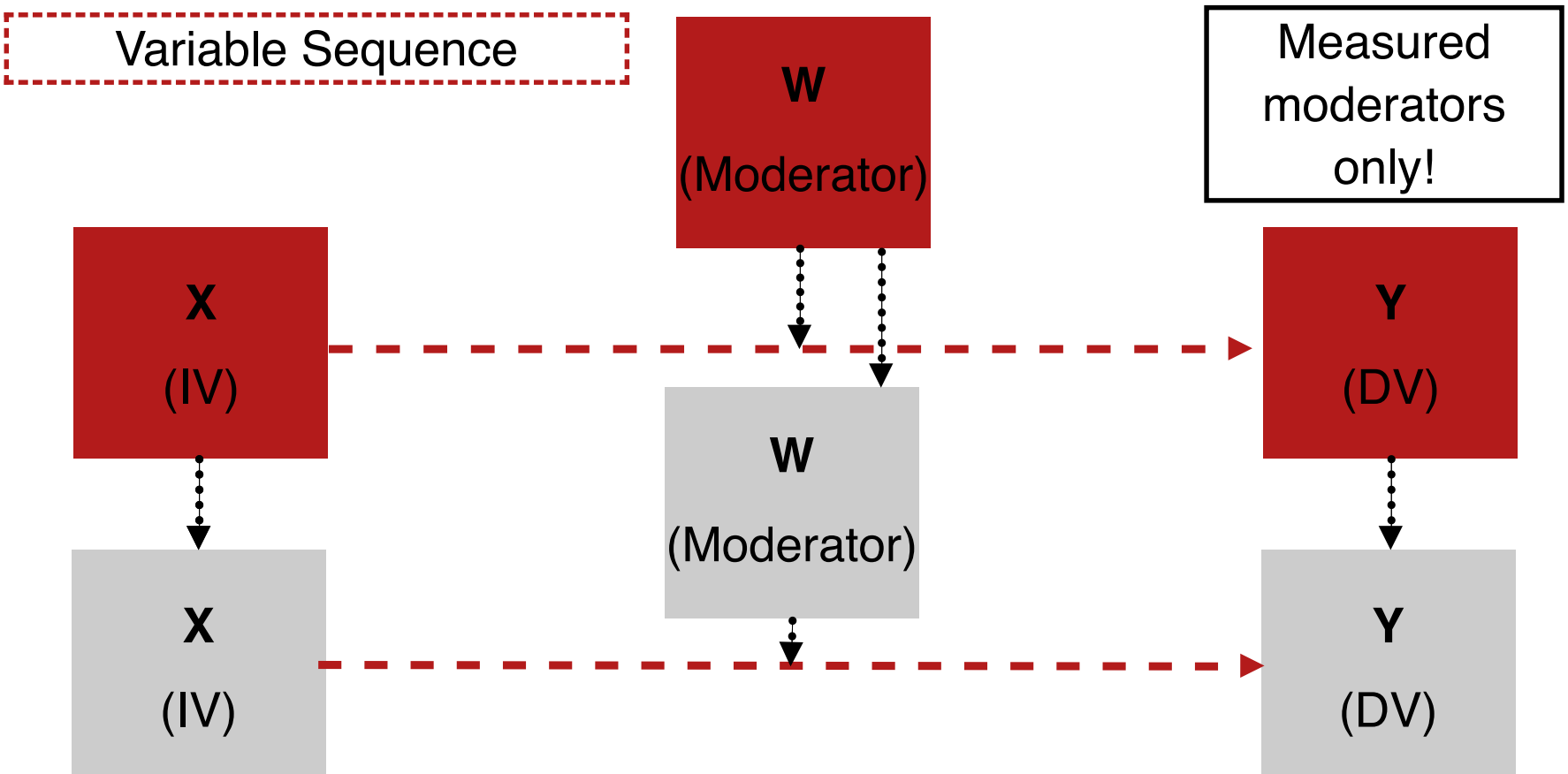
Moderation: What could go wrong?



Moderation: What could go wrong?



Moderation: What could go wrong?



Moderation: Best Practices

Manipulated Moderators

- Manipulated moderators have fewer validity threats, but need well-specified theory.
- Manipulated moderators should only manipulate the construct of interest.

Measured Moderators

- Subject to variable sequence threats (like mediation).
- Measured moderators should not be affected by other manipulations or measures.

All Moderation Designs

- Moderator cleanly maps to construct.
- Interaction alone doesn't always mean process evidence.
- Consider costs of recruitment.

I-FoMO and Investment Amount

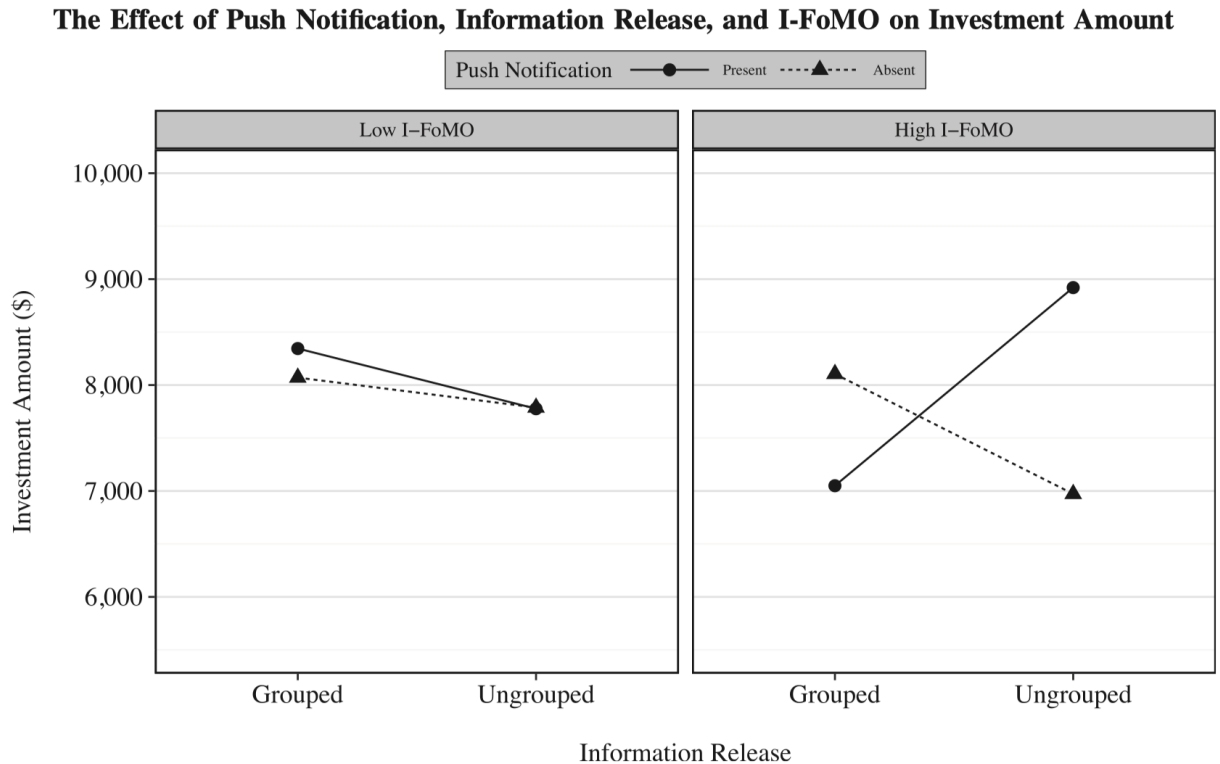


Figure 1 - Clor-Proell, Guggenmos, and Rennekamp (2020 TAR)

Multiple Sources

Increasing the quality, quantity, and/or diversity of process evidence can allow for stronger claims about process.

Multiple methods can help "rule out" that an effect is a methodological artifact.

Multiple experiments allow for strategic design.

- Breaking into smaller pieces (home in on process).
- Test the causal chain.
- Don't need to run cells that aren't useful.

Triangulation across studies can lead to richer inference over the lifetime of a research literature.

Trust in Algorithms

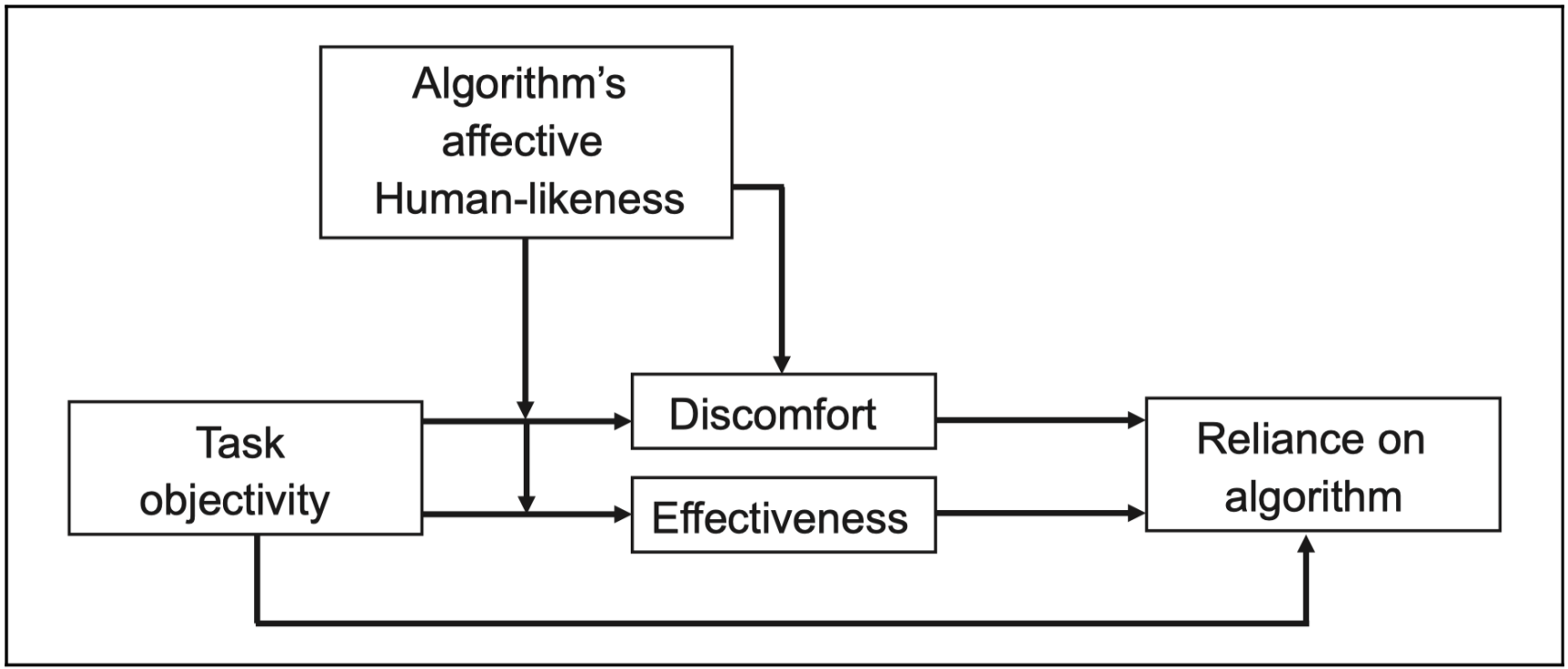


Figure 1 - Castelo, Bos, and Lehmann (2019 JMR)

Trust in Algorithms

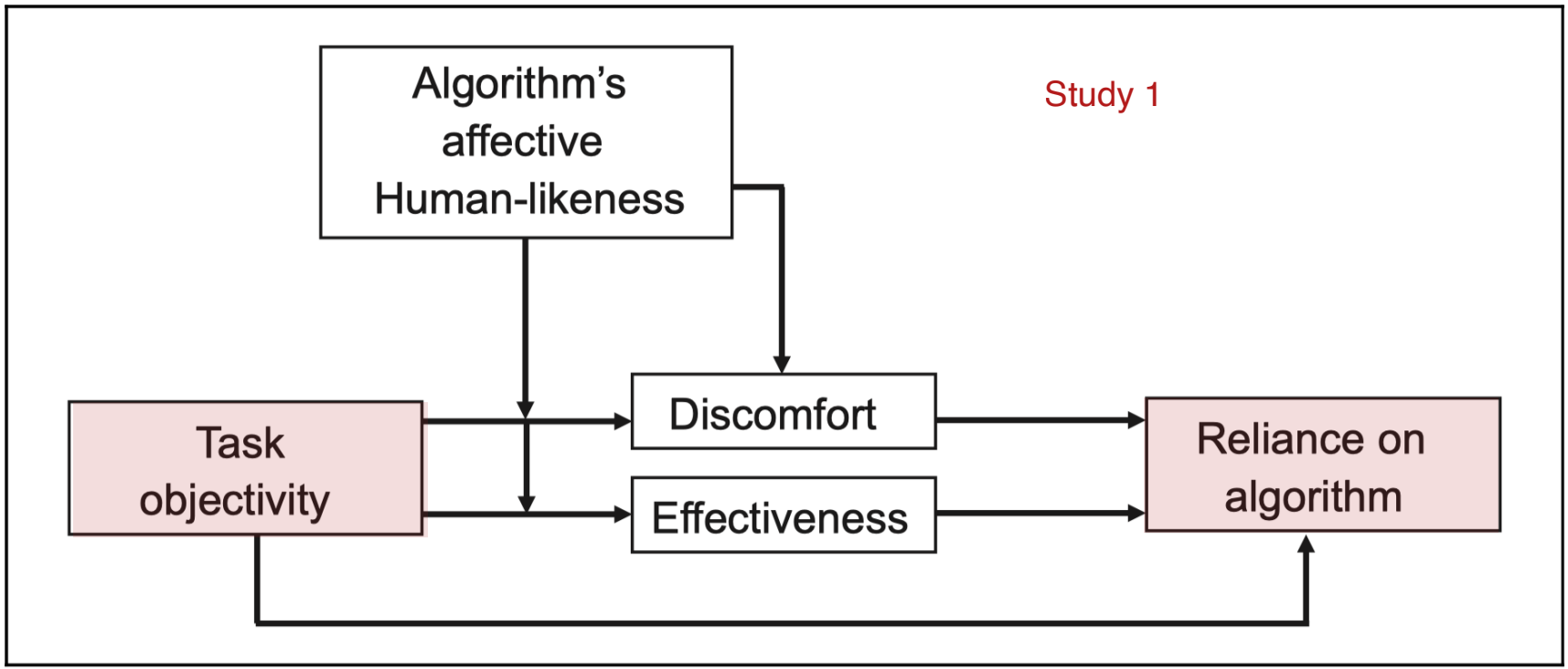


Figure 1 - Castelo, Bos, and Lehmann (2019 JMR)

Trust in Algorithms

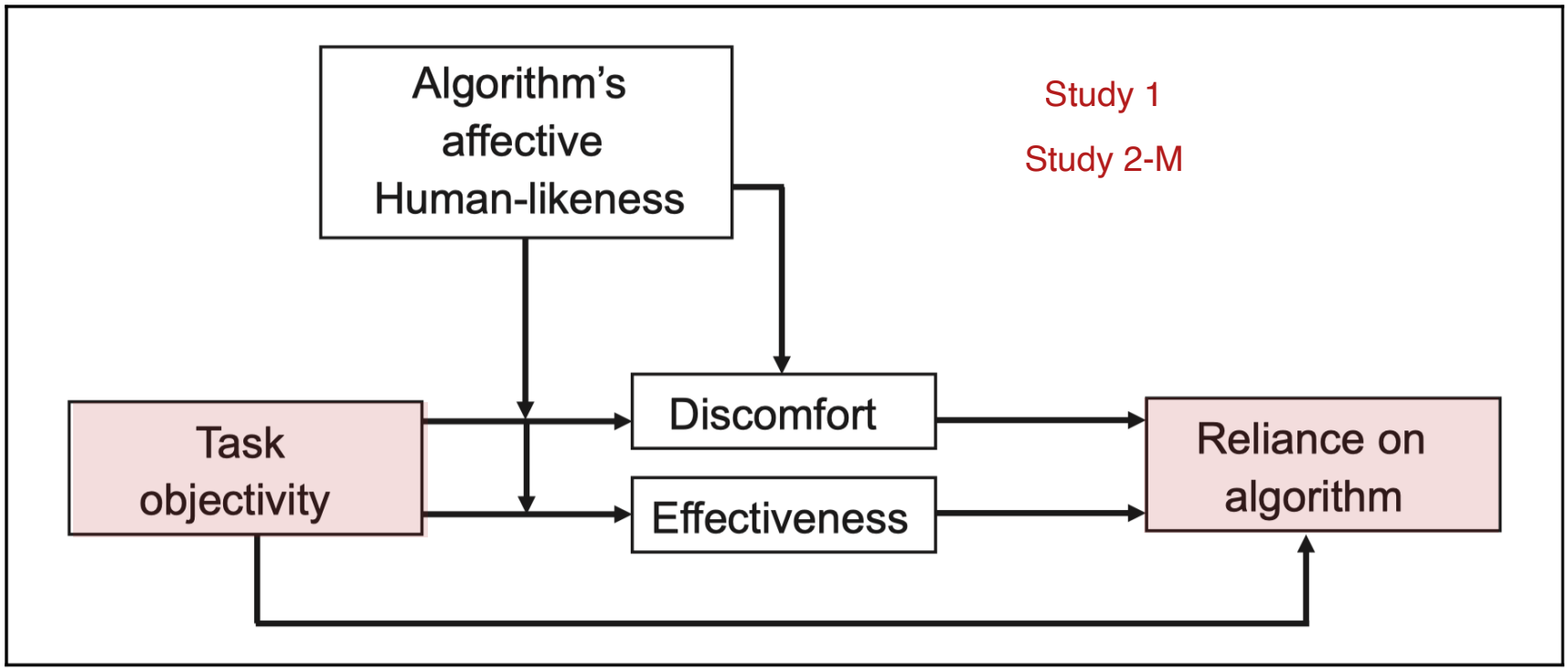


Figure 1 - Castelo, Bos, and Lehmann (2019 JMR)

Trust in Algorithms

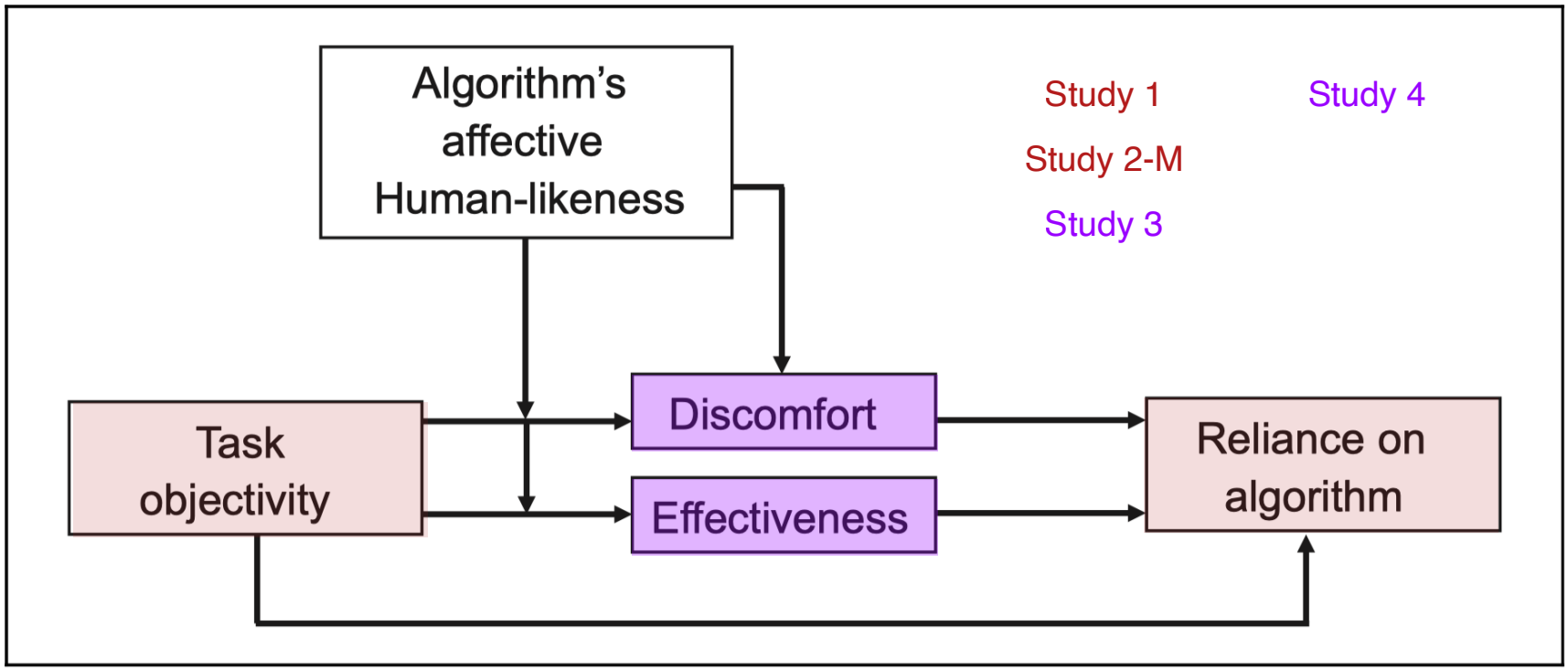


Figure 1 - Castelo, Bos, and Lehmann (2019 JMR)

Trust in Algorithms

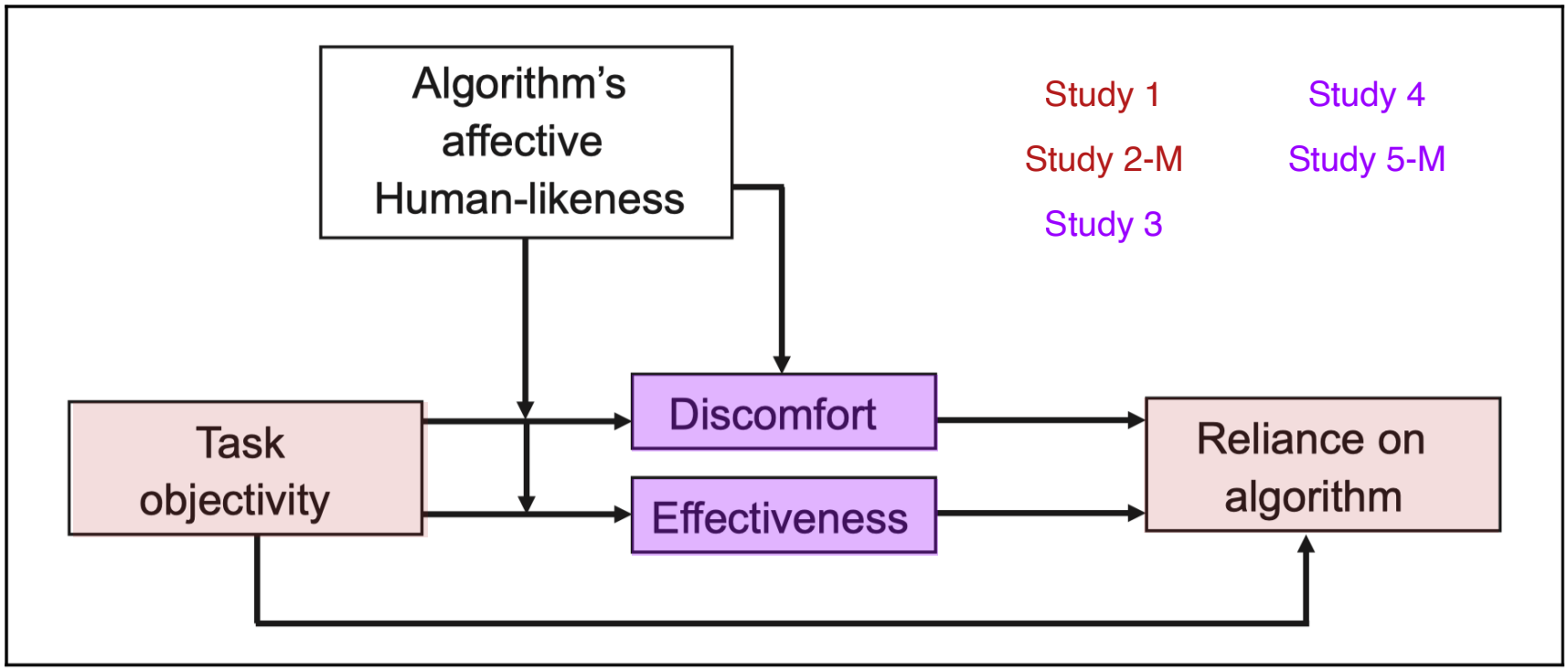


Figure 1 - Castelo, Bos, and Lehmann (2019 JMR)

Trust in Algorithms

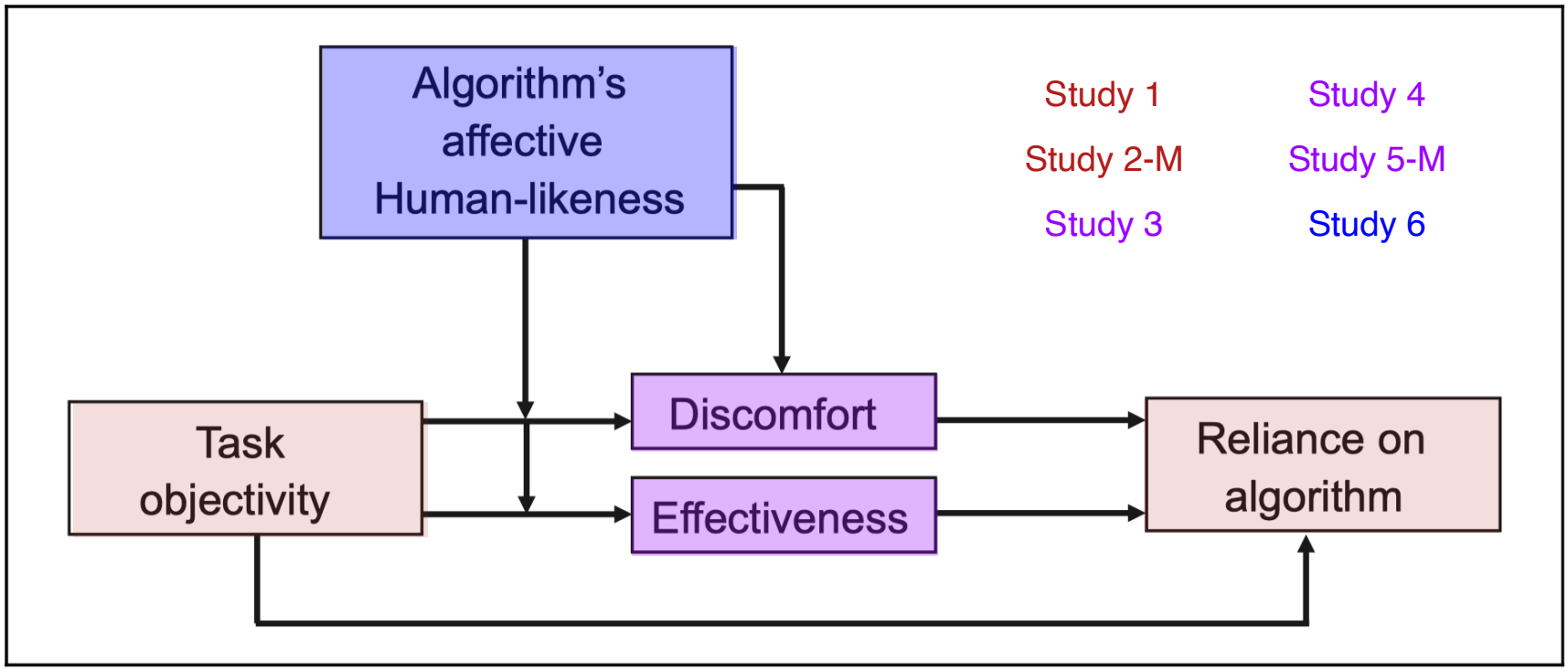


Figure 1 - Castelo, Bos, and Lehmann (2019 JMR)

Takeaways

Takeaways

Not every study needs to be able to tell us *how* or *why* something happens. Sometimes documenting an important effect is enough.

Takeaways

Not every study needs to be able to tell us *how* or *why* something happens. Sometimes documenting an important effect is enough.

But when we want to speak to process, approach providing process evidence with rigor commensurate with how much the contribution hinges on understanding process.

Takeaways

Not every study needs to be able to tell us *how* or *why* something happens. Sometimes documenting an important effect is enough.

But when we want to speak to process, approach providing process evidence with rigor commensurate with how much the contribution hinges on understanding process.

Process testing (and design, more generally) is a game of tradeoffs. Be mindful and make these choices intentionally.

Takeaways

Not every study needs to be able to tell us *how* or *why* something happens. Sometimes documenting an important effect is enough.

But when we want to speak to process, approach providing process evidence with rigor commensurate with how much the contribution hinges on understanding process.

Process testing (and design, more generally) is a game of tradeoffs. Be mindful and make these choices intentionally.

Choose the right tool for the job. Whether mediation vs. moderation, measurement vs. manipulation, or choosing between one or multiple studies - match the theory, to the hypothesis, to the method, to the analysis.

More Resources and References

For more resources related to statistics and research design:

https://rdg222.github.io/Resources/

Primary Reference:

Asay, H. S., R. Guggenmos, K. Kadous, L. Koonce, and R. Libby. 2021.“Theory Testing and Process Evidence in Accounting Experiments.” Working Paper (SSRN)

Thank you!