EMBRACE
Establishing Mechanisms of Benefit to Reinforce the Alzheimer's Care Experience

Examining Mechanisms of Intervention Impact using Statistical Mediation Methods

David L. Roth, PhD

Professor Emeritus
Johns Hopkins University

July 10, 2025

Disclosures

I have no disclosures to report.

Outline

1. Review the Basic Concepts and Terminology of Statistical Mediation Modeling
2. Applying Mediation Models to Multi-wave (e.g., pre/baseline, post 1, post 2) randomized trials of interventions.
3. Specific considerations
   A. Analyzing change scores (post – pre) vs. post scores only.
   B. Using baseline values as covariates, when available.
   C. Testing the statistical significance of the mediated effect.
   D. Examining mediation even when the intervention did not have a significant effect on the primary outcome.
   E. The biasing effects of unreliability in mediator measurement. Observed vs. Latent mediators.
   F. Modifications for when the primary outcome is binary, or a time-to-event outcome.
4. Examples of Applications

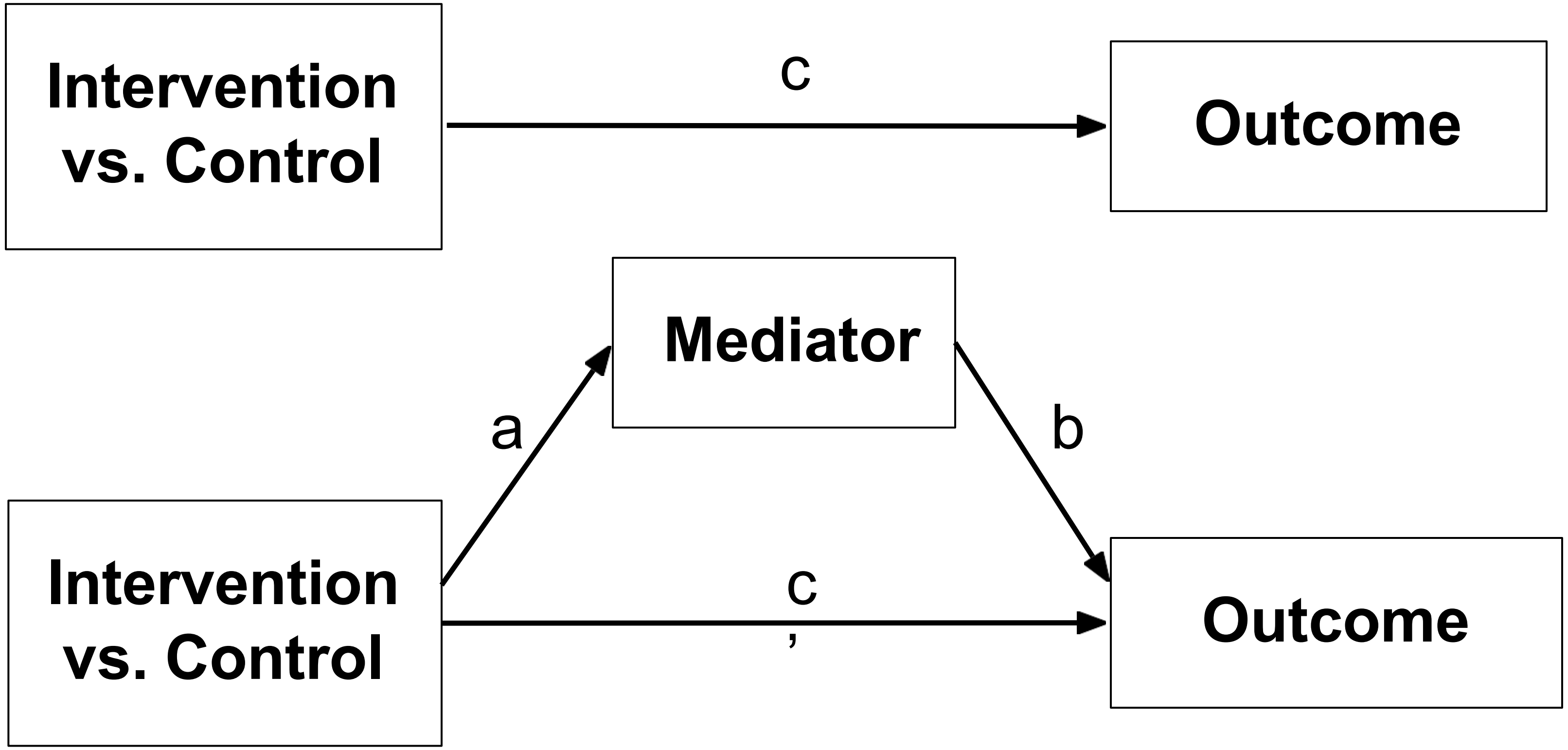
Learning Objectives

1. To understand the principles and concepts when testing mechanistic hypotheses with statistical mediation modeling.

2. To select from the various options involved including baseline covariate adjustments, change score analyses, data wave timing, and mediation measurement when designing mediation models.

3. To approximate the sample size needed to adequately power mediation testing.

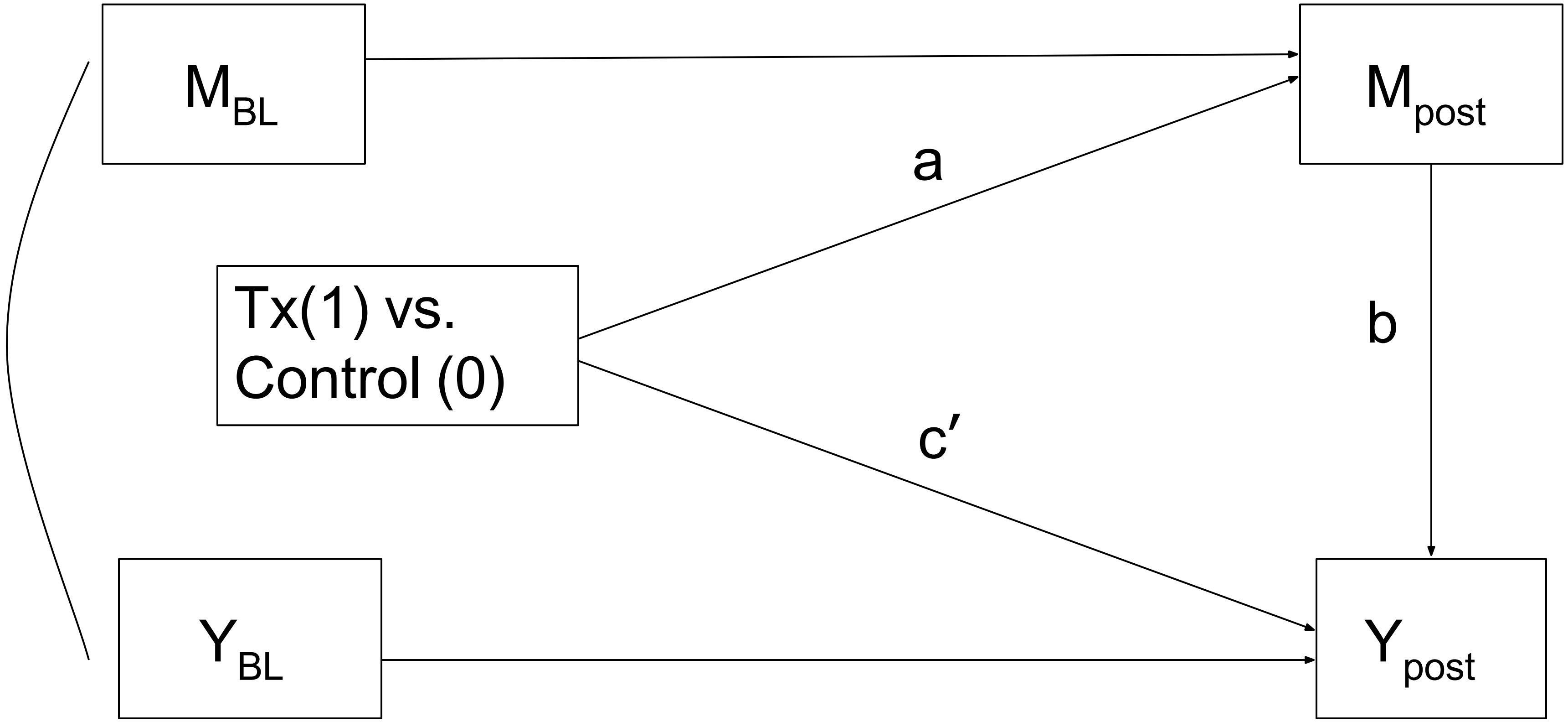
Basics of Statistical Mediation Modeling



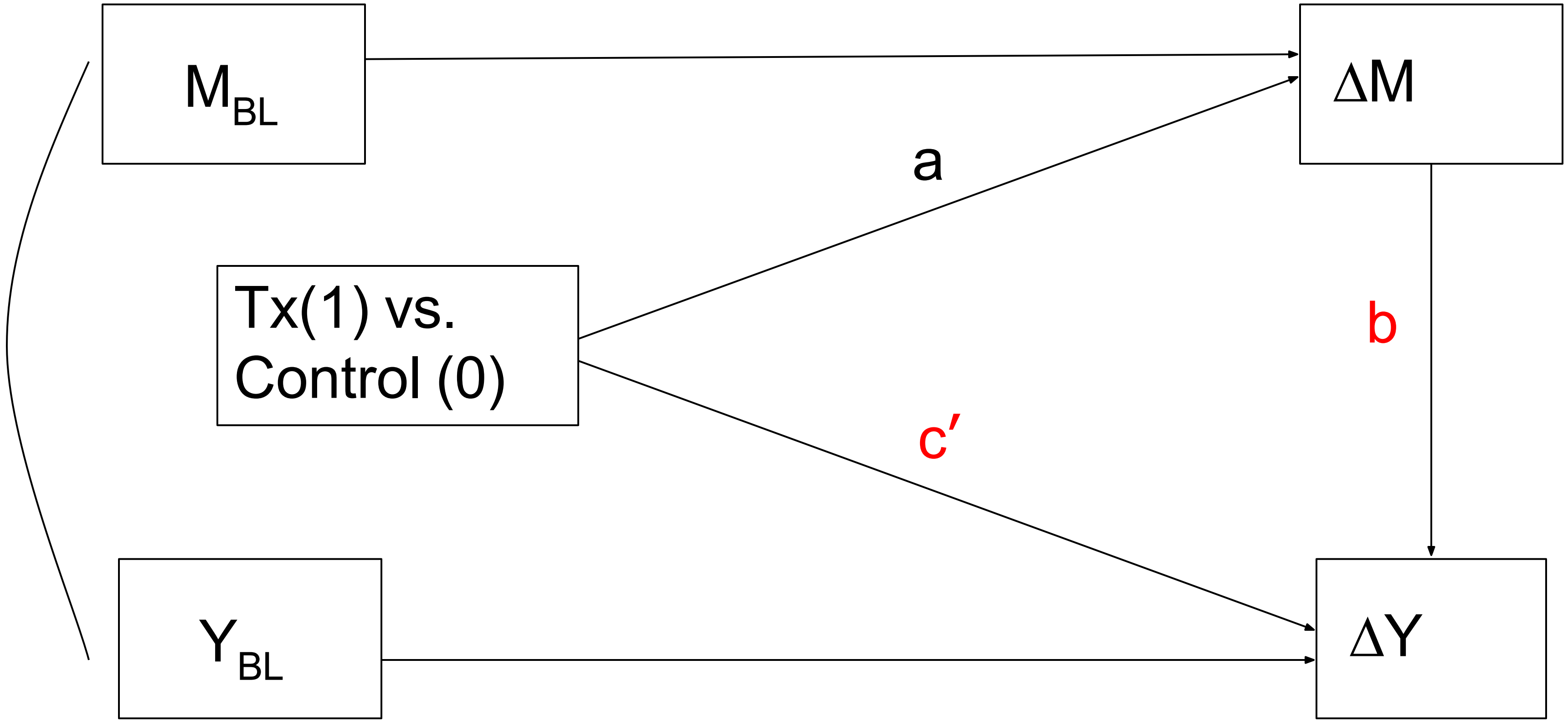
<u>Total effect</u>: c
<u>Direct (or unmediated) effect</u>: c'

<u>Mediated (or indirect) effect</u>: a*b, or c - c'
<u>Proportion mediated</u>: (a*b)/((a*b) + c'), or (c - c') / c

Mediation in a 2-wave (Pre-Post) RCT Design

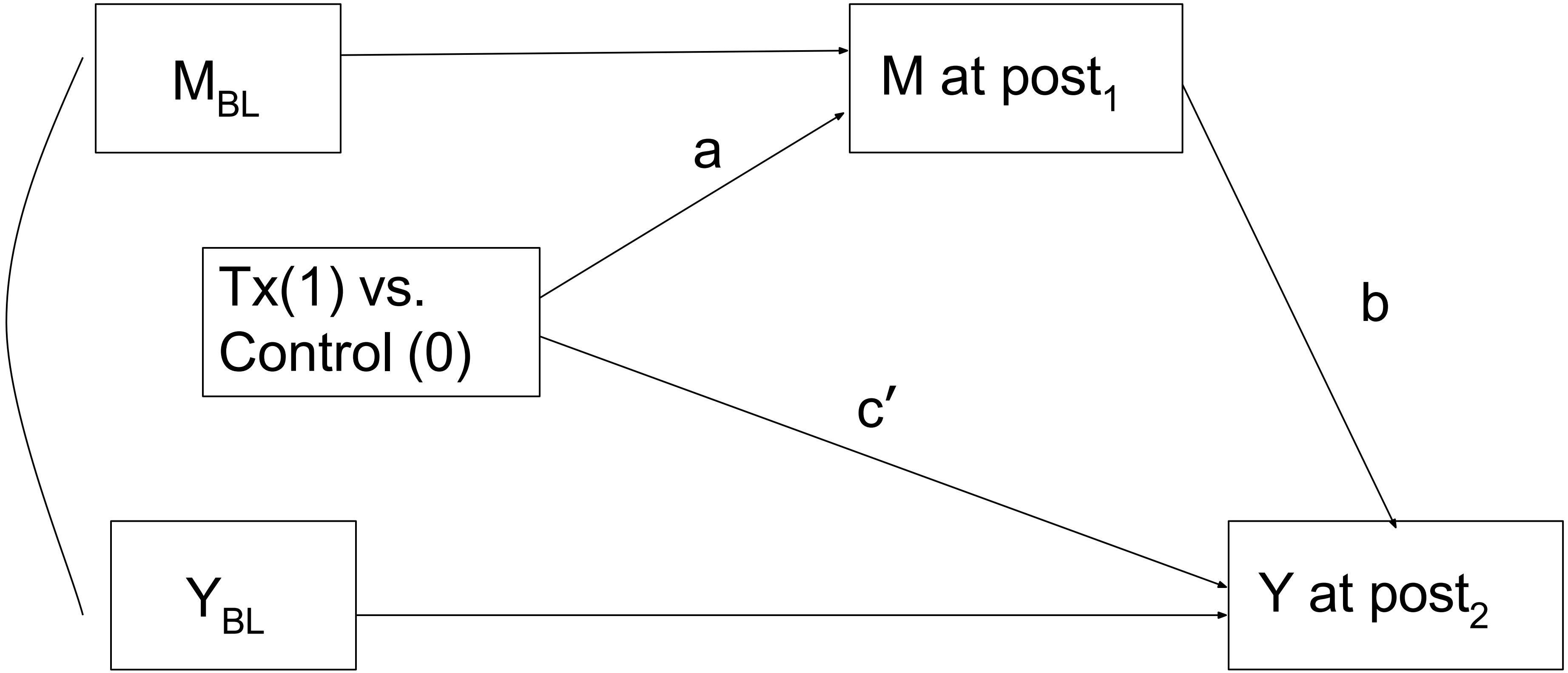


Mediation in a 2-wave (Pre-Post) RCT Design with Change Scores

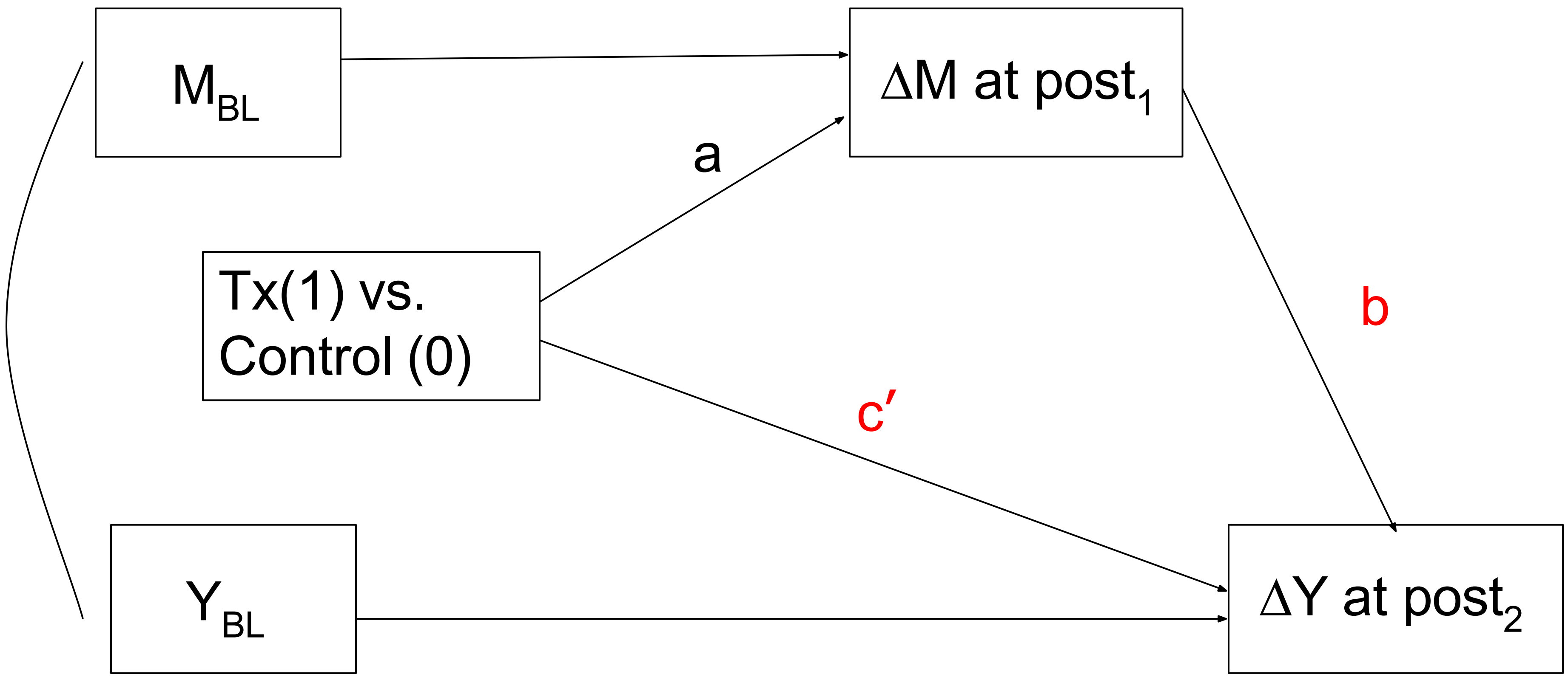


where Δ = post – pre change scores. Note that the b and c’ paths change from an analysis of post scores only.

Mediation in a 3-wave (Pre, $Post_1$, $Post_2$) RCT Design

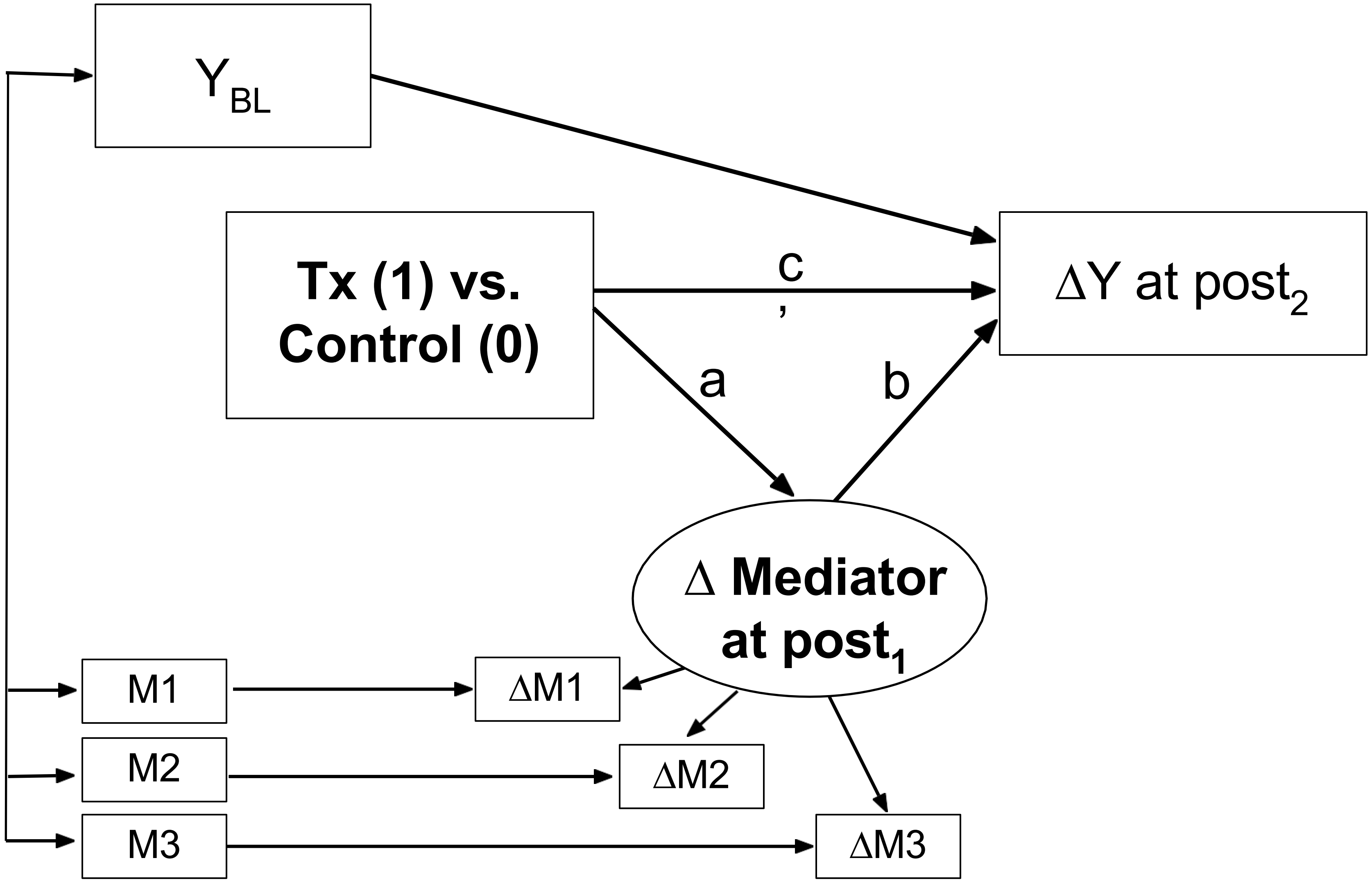


Mediation in a 3-wave (Pre, Post$_1$, Post$_2$) RCT Design



where Δ = post – pre change scores. Note that the b and c' paths change from an analysis of post scores only.

Mediation in a 3-wave (Pre, $Post_1$, $Post_2$) RCT Design with a Latent Mediator

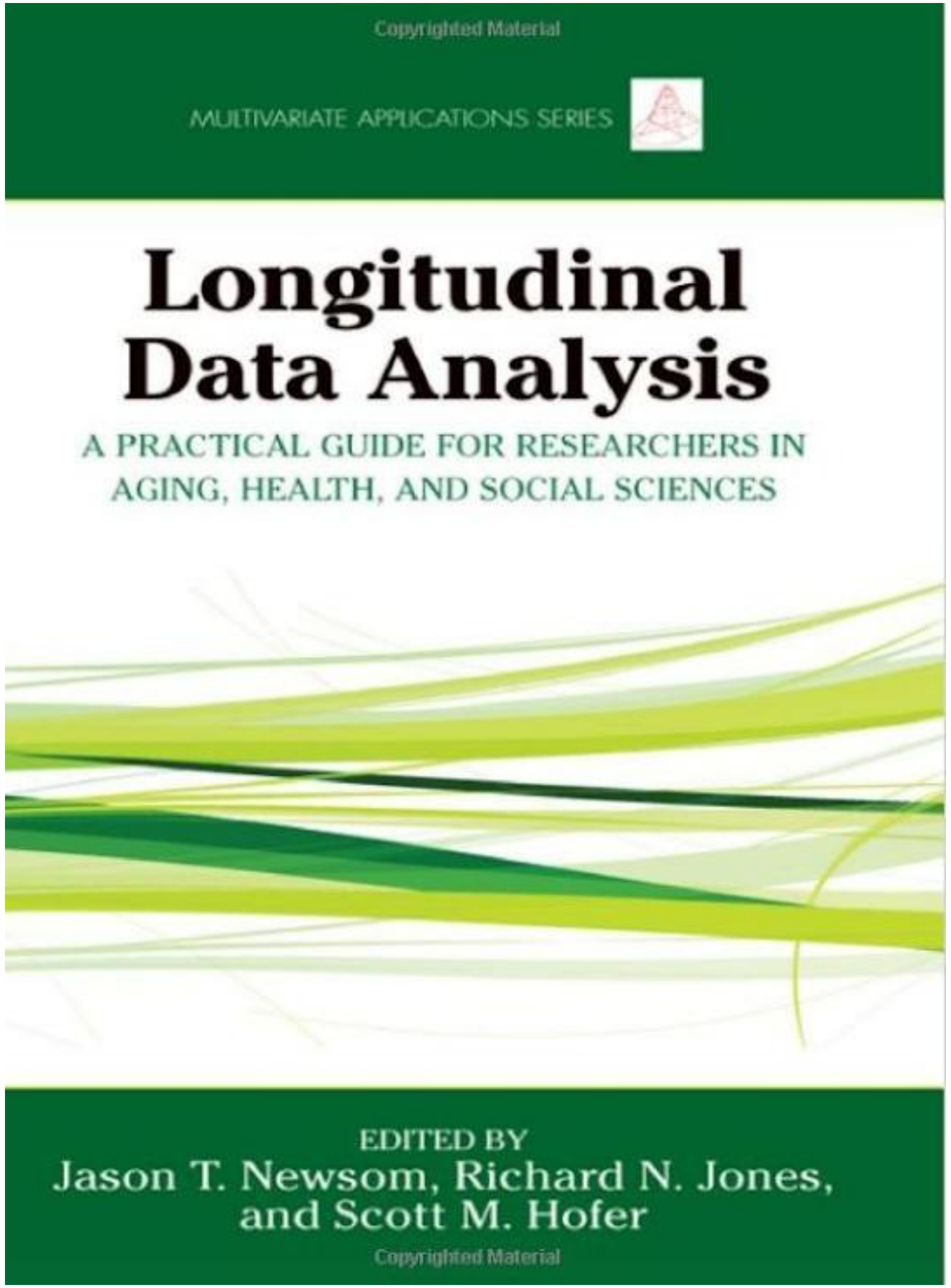


Testing the statistical significance of the mediated effect

1. Joint significance test. If both paths (a and b) of the mediated effect are statistically significant (e.g., $p < 0.05$), then the mediated effect is also statistically significant ($p < 0.05$).

2. Sobel (1982) test. Provides a formula to calculate the standard error for the a*b effect, then you divide a*b by this standard error and get a p-value from the standard normal distribution.

3. Bootstrapping methods. Resample the data, with replacement, thousands of times, calculate a*b from each sampling, then compute a standard error for a*b.


Copyrighted Material
MULTIVARIATE APPLICATIONS SERIES
Longitudinal Data Analysis
A PRACTICAL GUIDE FOR RESEARCHERS IN AGING, HEALTH, AND SOCIAL SCIENCES
EDITED BY
Jason T. Newsom, Richard N. Jones, and Scott M. Hofer
Copyrighted Material

Contents

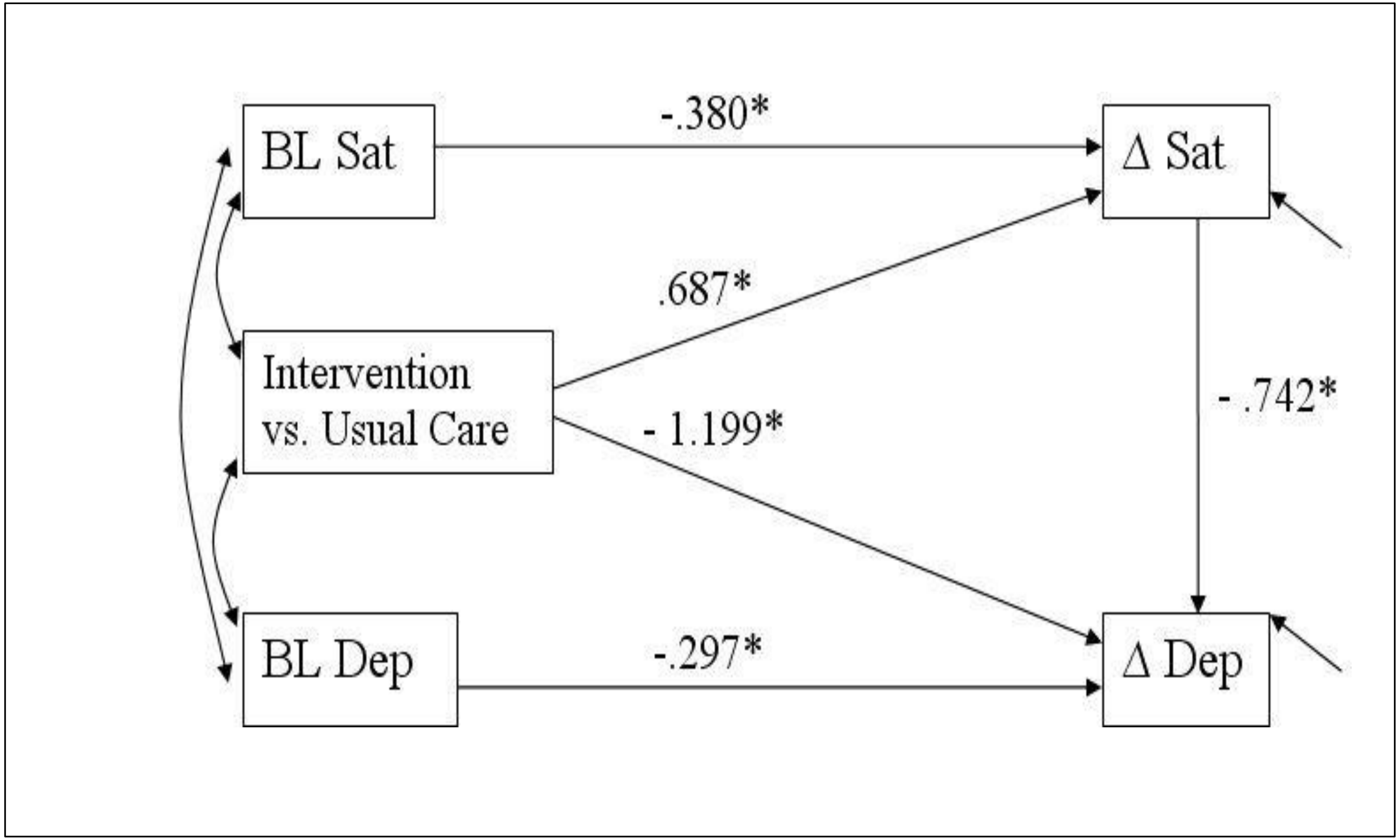
NYU Caregiver Intervention Study (M. Mittelman, PI)

- Randomized controlled trial of 406 spouse caregivers of dementia patients. Dyads were randomly assigned to an intervention condition or a usual care control group.
- The intervention involved strengthening *social support resources* (e.g., support group services, other family members). Usual care participants received information and access to standard services.
- 312 spouse caregivers provided care in the home for at least one year after randomization.
- Changes in social support were hypothesized to be important proximal outcomes and mediators of change in caregiver depression and patient nursing home placement rate.

Satisfaction with Social Support in the NYU Caregiver Intervention Study

Likert-type ratings were obtained on how satisfied spouse caregivers were with their social support networks (1 = very dissatisfied, 6 = very satisfied).

- “In ***general***, how satisfied are you with your social network?”
- “How satisfied are you with the ***assistance*** you get with daily activities (help with chores, patient care)?”
- “How satisfied are you with the ***emotional*** support you get from your social network?”


BL Sat
-.380*
Δ Sat
.687*
Intervention vs. Usual Care
- 1.199*
- .742*
BL Dep
-.297*
Δ Dep

% mediated = 30%

Psychology and Aging
2014, Vol. 29, No. 3, 601–611

0882-7974/14/$12.00 http://dx.doi.org/10.1037/a0036784

Mediators of the Impact of a Home-Based Intervention (Beat the Blues) on Depressive Symptoms Among Older African Americans

Laura N. Gitlin, David L. Roth, and Jin Huang
Johns Hopkins University

Older African Americans ($N = 208$) with depressive symptoms were randomly assigned to a home-based nonpharmacologic intervention (Beat the Blues, or BTB) or wait-list control group. BTB was delivered by licensed social workers and involved up to 10 home visits focused on care management, referral and linkage, depression knowledge and efficacy in symptom recognition, instruction in stress reduction techniques, and behavioral activation through identification of personal goals and action plans for achieving them. Structured interviews by assessors masked to study assignment were used to assess changes in depressive symptoms (main trial endpoint), behavioral activation, depression knowledge, formal care service utilization, and anxiety (mediators) at baseline and 4 months. At 4 months, the intervention had a positive effect on depressive symptoms and all mediators except formal care service utilization. Structural equation models indicated that increased activation, enhanced depression knowledge, and decreased anxiety each independently mediated a significant proportion of the intervention's impact on depressive symptoms as assessed with 2 different measures (PHQ-9 and CES-D). These 3 factors also jointly explained over 60% of the intervention's total effect on both indicators of depressive symptoms. Our findings suggest that most of the impact of BTB on depressive symptoms is driven by enhancing activation or becoming active, reducing anxiety, and improving depression knowledge/efficacy. The intervention components appear to work in concert and may be mutually necessary for maximal benefits from treatment to occur. Implications for designing tailored interventions to address depressive symptoms among older African Americans are discussed.

Keywords: depression, mediation models, mental health disparities

Beat the Blues Intervention (L. Gitlin, PI)

- Randomized controlled trial of 208 African Americans age 55+ with PHQ-9 scores $\geq$ 5. Participants were randomly assigned to a multicomponent cognitive-behavioral intervention or to a wait-list control.
- Beat the Blues was delivered by social workers and targeted symptom recognition, depression knowledge, stress reduction, and behavioral activation.
- 179 participants provided mediator and outcome data at 4 weeks. Changes in depression knowledge, behavioral activation, and anxiety
- hypothesized to mediate the effect of the intervention on changes in depressive symptoms as measured by two outcomes (PHQ-9 and CES-D).
- Mediators were examined individually and jointly.

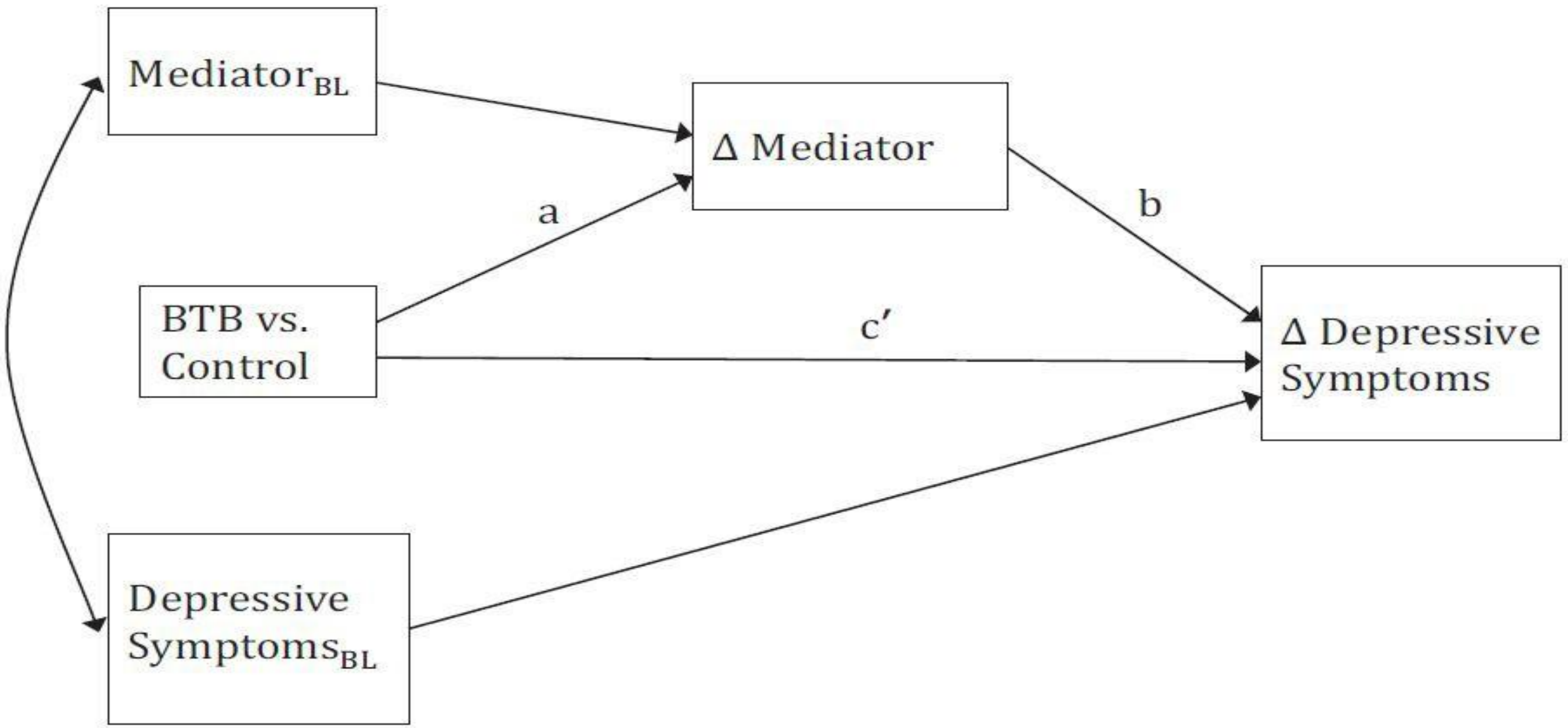


Figure 1. Two-wave mediation model used to examine mediators individually. BL = baseline observation. Δ = 4-month score minus baseline score.

Table 3

Summary of Single Mediator Models on Depression Measures

Outcome	Mediator	a	b	ab	c'	ab/(ab + c')
PHQ9	Behavioral activation	0.71^{***}	-1.21^{***}	-0.86^{**}	-1.74^{*}	0.33
	Depression knowledge	0.31^{***}	-3.20^{***}	-0.99^{**}	-1.46	0.41
	Anxiety	-0.33^{***}	2.69^{***}	-0.89^{**}	-1.67^{*}	0.35
	Formal care	0.10	-0.07	-0.01	-2.53^{**}	0.00
CESD	Behavioral activation	0.71^{***}	-1.82^{***}	-1.29^{***}	-1.93^{**}	0.40
	Depression knowledge	0.31^{***}	-3.83^{***}	-1.19^{***}	-1.84^{*}	0.39
	Anxiety	-0.33^{***}	3.92^{***}	-1.30^{**}	-1.87^{**}	0.41
	Formal care	0.10	0.00	0.00	-3.12^{***}	0.00

Note. PHQ-9 = Patient Health Questionnaire; CES-D = Center for Epidemiologic Studies–Depression scale; Refer to Figure 1 to understand letters, a, b, ab, c', ab/(ab + c'), at top of columns.
$^{*}p < 0.05$. $^{**}p < 0.01$. $^{***}p < 0.001$.

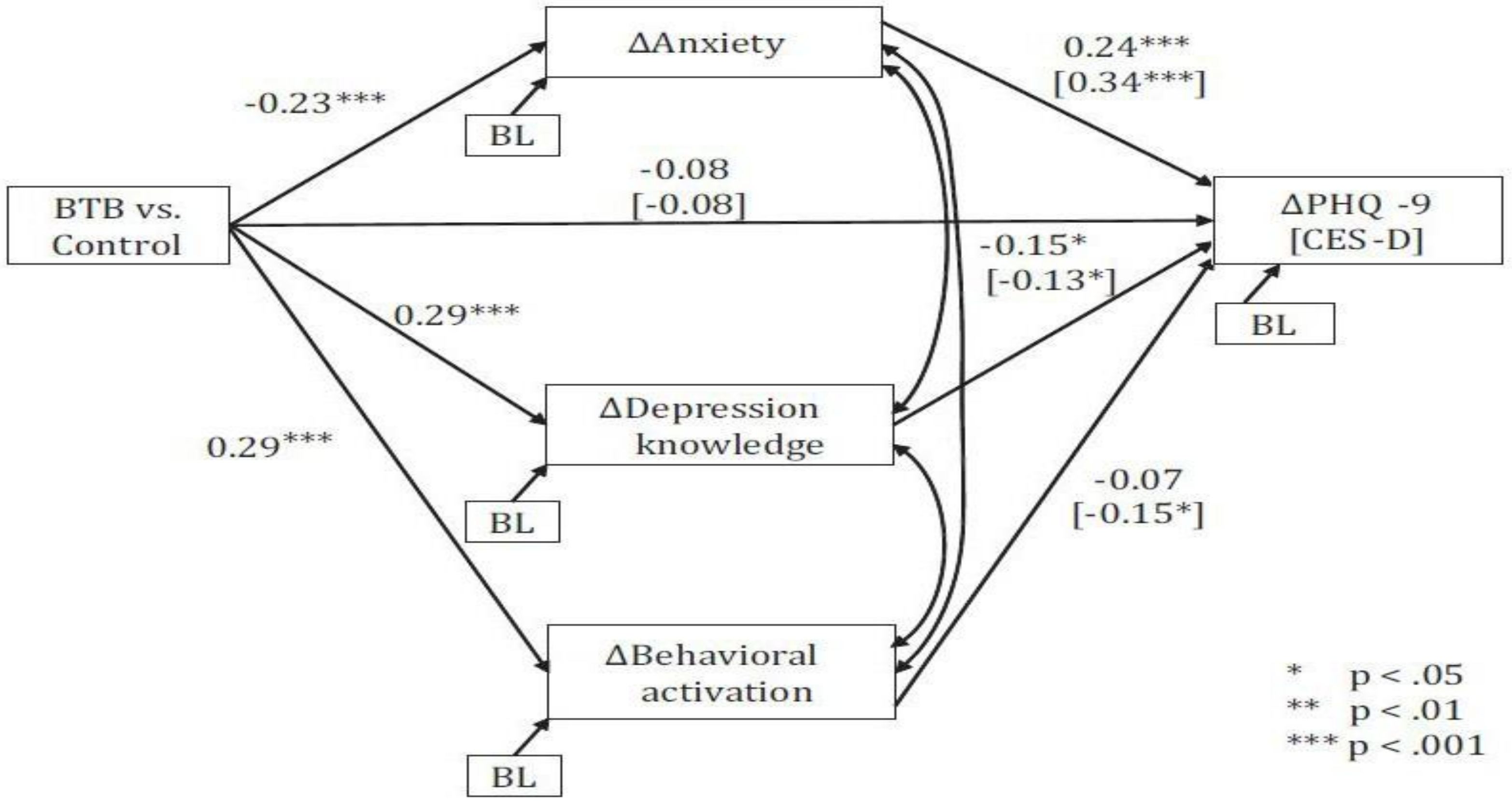


Figure 2. Multiple-mediator model of intervention effect on change in depressive symptoms. BL = baseline observation. Δ = 4-month score minus baseline score. PHQ-9 = Patient Health Questionnaire; CES-D = Center for Epidemiologic Studies–Depression scale. Numbers on top refer

Psychology and Aging

ISSN: 0882-7974

2025, Vol. 40, No. 4, 355–370
https://doi.org/10.1037/pag0000893

Testing the Purported Mechanisms of the AgingPLUS Intervention: Effects on Physical Activity Outcomes

Manfred Diehl[1], Han-Yun Tseng[1], George W. Rebok[2, 3], Kaigang Li[4], Abigail M. Nehrkorn-Bailey[5], Diana Rodriguez[1], Diefei Chen[2], and David L. Roth[2]
[1] Department of Human Development and Family Studies, Colorado State University
[2] Center on Aging and Health, School of Medicine, Johns Hopkins University
[3] Department of Mental Health, Bloomberg School of Public Health, Johns Hopkins University
[4] Department of Health and Exercise Science, Colorado State University
[5] Department of Psychology, University of Wisconsin-Green Bay

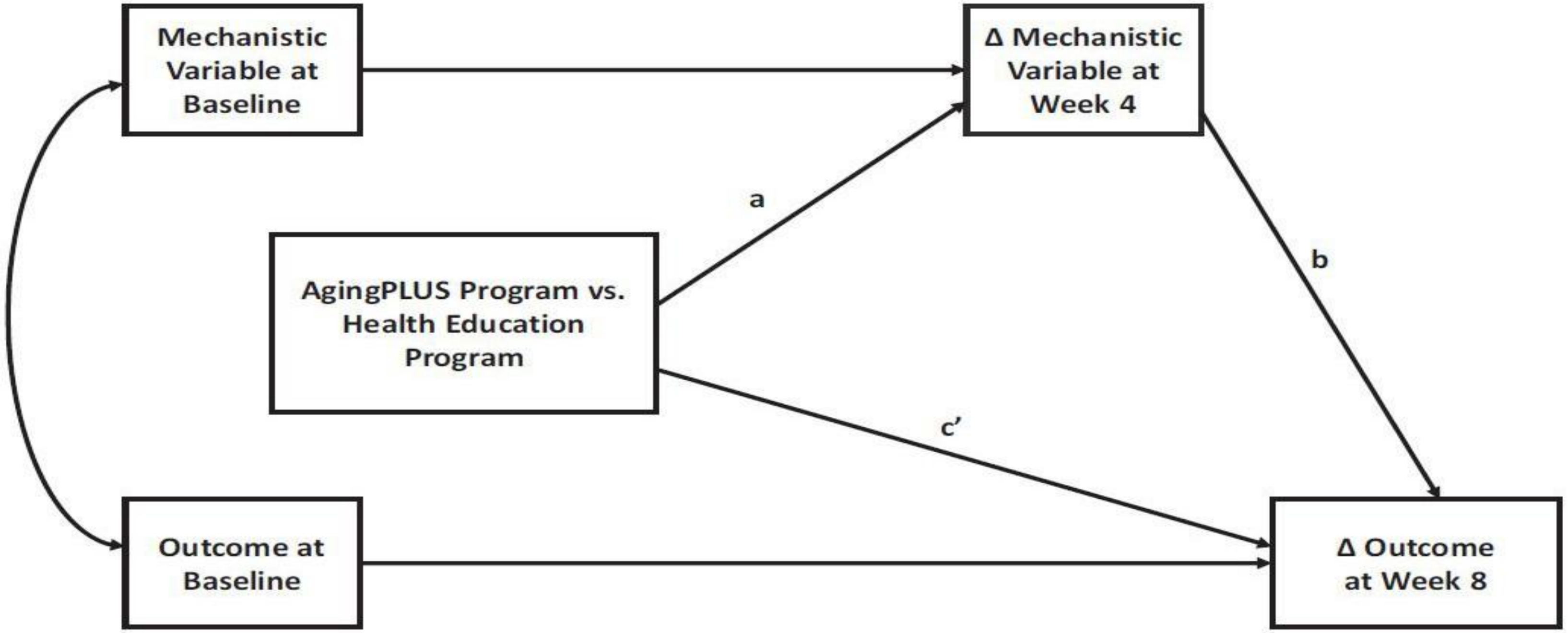
Following the experimental medicine approach, Diehl et al. (2023) demonstrated the malleability of negative views of aging (NVOA), self-efficacy beliefs, and exercise intention in middle-aged and older adults who participated in the AgingPLUS intervention program. The present study built on those findings and addressed (a) whether the intervention resulted in significant improvements in physical activity (PA) and (b) whether the purported mechanistic variables were significant mediators of the intervention's effects on PA outcomes. AgingPLUS used a randomized, single-blind control group design to implement the intervention in a sample of 335 adults aged 45–75 years. This study reports findings from 278 participants

EMBRACE

AgingPLUS Intervention (M. Diehl, PI)

- Randomized controlled trial of 335 adults 45-75 years of age. Participants were randomly assigned to the AgingPLUS intervention or to a Health Education control condition.
- AgingPLUS aimed to improve physical activity by targeting age-related motivational factors. These included negative views of aging, self-efficacy, and behavioral intentions.
- 278 participants provided physical activity outcome data at 8 weeks.
- Changes in motivational factors measured at 4 weeks were hypothesized to mediate the effect of the intervention on changes in physical activity at 8 weeks.
- 108 separate mediation models were run, 9 mediators for 12 physical activity outcomes (9*12 = 108).

Figure 2
Conceptual Model Testing the Mechanistic Assumptions of the Intervention



Note. a = the association between the intervention and the change in the mechanistic variable at Week 4; b = the association between the change in the mechanistic variable at Week 4 and the change in the outcome variable at Week 8; and c' = the association between the intervention program and the change in the outcome variable at Week 8.

Table 4
Findings From Mediation Analyses Supporting Significant Effects of the Purported Mechanistic Variables

Outcome variable at Week 8	Mediator at Week 4	Direct effect (c')	Path a	Path b	Bootstrapped indirect effect ($a \times b$)	Bootstrapped 95% CI	
						LL	UL
Accelerometer (per day)							
Total kcals burned	AS	$0.12^{\dagger}$	0.34^{***}	0.17^{**}	0.06^{**}	0.01	0.11
Total kcals burned	ERA	0.10	0.34^{***}	0.21^{**}	0.07^{**}	0.03	0.12
Total minutes of light PA	AS	−0.06	0.34^{***}	0.23^{***}	0.08^{**}	0.04	0.13
% of total minutes of light PA	AS	−0.09	0.34^{***}	0.23^{***}	0.08^{**}	0.03	0.12
Total minutes of MVPA	ERA	0.14^{*}	0.34^{***}	0.15^{*}	0.05^{*}	0.01	0.10
% of total minutes of MVPA	ERA	$0.11^{\dagger}$	0.34^{***}	0.15^{*}	0.05^{*}	0.01	0.10
% of total minutes of MVPA	GSE	0.14^{*}	0.17^{**}	0.14^{*}	0.02^{*}	0.002	0.06
CHAMPS (per week)							
Frequency of MVPA	AS	0.02	0.36^{***}	0.14^{*}	0.06^{**}	0.02	0.11
Frequency of MVPA	GSE	0.04	0.17^{***}	0.12^{*}	0.02^{*}	0.002	0.06
Frequency of MVPA	MSE	0.04	0.25^{***}	0.14^{*}	0.04^{*}	0.01	0.07
Frequency of MVPA	VSE	0.04	0.22^{***}	0.15^{**}	0.04^{*}	0.01	0.07
DAL (per week)							
Total minutes of light PA	ERA	−0.08	0.36^{***}	-0.10^{*}	-0.03^{*}	−0.07	−0.01
Total minutes of MVPA	AS	−0.01	0.35^{***}	0.19^{***}	0.07^{**}	0.02	0.16
Total minutes of MVPA	ERA	0.02	0.36^{***}	0.11^{*}	0.07^{**}	0.01	0.08
Total minutes of MVPA	EBA	0.02	0.22^{***}	0.16^{**}	0.04^{*}	0.01	0.09
Total minutes of MVPA	GSE	0.03	0.17^{**}	0.12^{*}	0.02^{*}	0.01	0.05

Note. Standardized coefficients are reported. PA = physical activity; MVPA = moderate-to-vigorous physical activity; AS = age stereotypes; ERA = expectations regarding aging; GSE = general self-efficacy; MSE = motivational self-efficacy; VSE = volitional self-efficacy; EBA = essential beliefs about aging; CI = confidence interval; *LL* = lower limit; *UL* = upper limit; DAL = daily activity log; kcals = kilocalories; CHAMPS = Community Healthy Activities Model Program for Seniors.
$^{\dagger}p < .10$. $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

Table 4
Findings From Mediation Analyses Supporting Significant Effects of the Purported Mechanistic Variables

Outcome variable at Week 8	Mediator at Week 4	Direct effect (c')	Path a	Path b	Bootstrapped indirect effect ($a \times b$)	Bootstrapped 95% CI	
						LL	UL
Accelerometer (per day)							
Total kcals burned	AS	$0.12^{\dagger}$	0.34^{***}	0.17^{**}	0.06^{**}	0.01	0.11
Total kcals burned	ERA	0.10	0.34^{***}	0.21^{**}	0.07^{**}	0.03	0.12
Total minutes of light PA	AS	−0.06	0.34^{***}	0.23^{***}	0.08^{**}	0.04	0.13
% of total minutes of light PA	AS	−0.09	0.34^{***}	0.23^{***}	0.08^{**}	0.03	0.12
Total minutes of MVPA	ERA	0.14^{*}	0.34^{***}	0.15^{*}	0.05^{*}	0.01	0.10
% of total minutes of MVPA	ERA	$0.11^{\dagger}$	0.34^{***}	0.15^{*}	0.05^{*}	0.01	0.10
% of total minutes of MVPA	GSE	0.14^{*}	0.17^{**}	0.14^{*}	0.02^{*}	0.002	0.06
CHAMPS (per week)							
Frequency of MVPA	AS	0.02	0.36^{***}	0.14^{*}	0.06^{**}	0.02	0.11
Frequency of MVPA	GSE	0.04	0.17^{***}	0.12^{*}	0.02^{*}	0.002	0.06
Frequency of MVPA	MSE	0.04	0.25^{***}	0.14^{*}	0.04^{*}	0.01	0.07
Frequency of MVPA	VSE	0.04	0.22^{***}	0.15^{**}	0.04^{*}	0.01	0.07
DAL (per week)							
Total minutes of light PA	ERA	−0.08	0.36^{***}	-0.10^{*}	-0.03^{*}	−0.07	−0.01
Total minutes of MVPA	AS	−0.01	0.35^{***}	0.19^{***}	0.07^{**}	0.02	0.16
Total minutes of MVPA	ERA	0.02	0.36^{***}	0.11^{*}	0.07^{**}	0.01	0.08
Total minutes of MVPA	EBA	0.02	0.22^{***}	0.16^{**}	0.04^{*}	0.01	0.09
Total minutes of MVPA	GSE	0.03	0.17^{**}	0.12^{*}	0.02^{*}	0.01	0.05

Note. Standardized coefficients are reported. PA = physical activity; MVPA = moderate-to-vigorous physical activity; AS = age stereotypes; ERA = expectations regarding aging; GSE = general self-efficacy; MSE = motivational self-efficacy; VSE = volitional self-efficacy; EBA = essential beliefs about aging; CI = confidence interval; LL = lower limit; UL = upper limit; DAL = daily activity log; kcals = kilocalories; CHAMPS = Community Healthy Activities Model Program for Seniors.
$^{\dagger}p < .10$. $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

Mplus Code for Diehl et al. (2025) mediation analysis

```
Title: group(X)- AS(M)-actigraph:kcal(Y);

Data: File is AgingPLUS_PA_Final_1206.csv;

Variable:
   Names are group d_kcal w0kcal w0AS d_AS ...
   (list all variables in *.csv file);
   Usevariables are group d_kcal w0kcal w0AS d_AS;
   Missing = .;

Analysis:
   bootstrap = 10000;

Model:
   d_AS ON w0AS group;
   d_kcal ON w0kcal d_AS group;

   model indirect:
   d_kcal IND group;

Output: stdyx cint(bcbootstrap);
```

Select Mplus output for Diehl et al. (2025) mediation analysis

STANDARDIZED MODEL RESULTS

STDYX Standardization

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value	
D_AS ON					
W0AS	-0.486	0.051	-9.627	0.000	
GROUP	0.339	0.048	7.046	0.000	a path
D_KCAL ON					
W0KCAL	-0.367	0.094	-3.885	0.000	
D_AS	0.167	0.070	2.402	0.016	b path
GROUP	0.118	0.057	2.060	0.039	c' path

Effects from GROUP to D_KCAL

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
Total	0.174	0.055	3.192	0.001
Total indirect	0.057	0.025	2.247	0.025

Bootstrapped standard error

PSYCHOLOGICAL SCIENCE

Research Article

Required Sample Size to Detect the Mediated Effect

Matthew S. Fritz and David P. MacKinnon

Arizona State University

ABSTRACT—*Mediation models are widely used, and there are many tests of the mediated effect. One of the most common questions that researchers have when planning mediation studies is, "How many subjects do I need to achieve adequate power when testing for mediation?" This article presents the necessary sample sizes for six of the most common and the most recommended tests of mediation for various combinations of parameters, to provide a guide for researchers when designing studies or applying for grants.*

Since the publication of Baron and Kenny's (1986) article describing a method to evaluate mediation, the use of mediation models in the social sciences has increased dramatically. Using the *Social Science Citation Index*, MacKinnon, Lockwood, Hoffman, West, and Sheets (2002) found more than 2,000 ci-

MEDIATION

In a mediation model, the effect of an independent variable (X) on a dependent variable (Y) is transmitted through a third intervening, or mediating, variable (M). That is, X causes M, and M causes Y. Figure 1 shows the path diagrams for a simple mediation model; the top diagram represents the total effect of X on Y, and the bottom diagram represents the indirect effect of X on Y through M and the direct effect of X on Y controlling for M. If M is held constant in a model in which the mediator explains all of the variation between X and Y (i.e., a model in which there is complete mediation), then the relationship between X and Y is zero.

The path diagrams in Figure 1 can be expressed in the form of three regression equations:

$$\hat{Y} = \hat{\zeta}_1 + \hat{\tau} X \qquad (1)$$

dence intervals detect the mediated effect.

RESULTS

Complete results are shown in Table 3. The sample sizes necessary to achieve .8 power in Baron and Kenny's (1986) test were very large for all of the complete-mediation ($\tau' = 0$) conditions

yielded identical results for the different τ' conditions, and results are therefore collapsed across these conditions in the table.

DISCUSSION

The most important result from this study is the finding that for Baron and Kenny's (1986) test, a sample size of at least 20,886

TABLE 3

Empirical Estimates of Sample Sizes Needed for .8 Power

	Condition															
Test	SS	SH	SM	SL	HS	HH	HM	HL	MS	MH	MM	ML	LS	LH	LM	LL
BK ($\tau' = 0$)	20,886	6,323	3,039	1,561	6,070	1,830	883	445	2,682	820	397	204	1,184	364	175	92
BK ($\tau' = .14$)	562	445	427	414	444	224	179	153	425	178	118	88	411	147	84	53
BK ($\tau' = .39$)	531	403	402	403	405	158	124	119	405	125	75	59	405	122	60	38
BK ($\tau' = .59$)	530	404	402	403	406	158	124	120	405	125	74	58	404	122	59	36
Joint significance	530	402	403	403	407	159	124	120	405	125	74	58	405	122	59	36
Sobel	667	450	422	412	450	196	144	127	421	145	90	66	410	129	67	42
PRODCLIN	539	402	401	402	402	161	125	120	404	124	74	57	404	121	58	35
Percentile bootstrap	558	412	406	398	414	162	126	122	404	124	78	59	401	123	59	36
Bias-corrected bootstrap	462	377	400	385	368	148	115	118	391	116	71	53	396	115	54	34

Note. All sample sizes have been rounded up to the next whole number. In the condition labels, the first letter refers to the size of the α path, and the second letter refers to the size of the β path; S = 0.14, H = 0.26, M = 0.39, and L = 0.59 (e.g., condition SM is the condition with $\alpha = 0.14$ and $\beta = 0.39$). All results, except for those for Baron and Kenny's (1986) test (BK), have been collapsed across τ' conditions.

NYU Caregiver Intervention Study (M. Mittelman, PI)

- Randomized controlled trial of 406 spouse caregivers of dementia patients. Dyads were randomly assigned to an intervention condition or a usual care control group.
- The intervention involved strengthening ***social support resources*** (e.g., support group services, other family members). Usual care participants received information and access to standard services.

- 312 spouse caregivers provided care in the home for at least one year after randomization.
- Changes in social support were hypothesized to be important proximal outcomes and mediators of change in caregiver depression and patient nursing home placement rate.

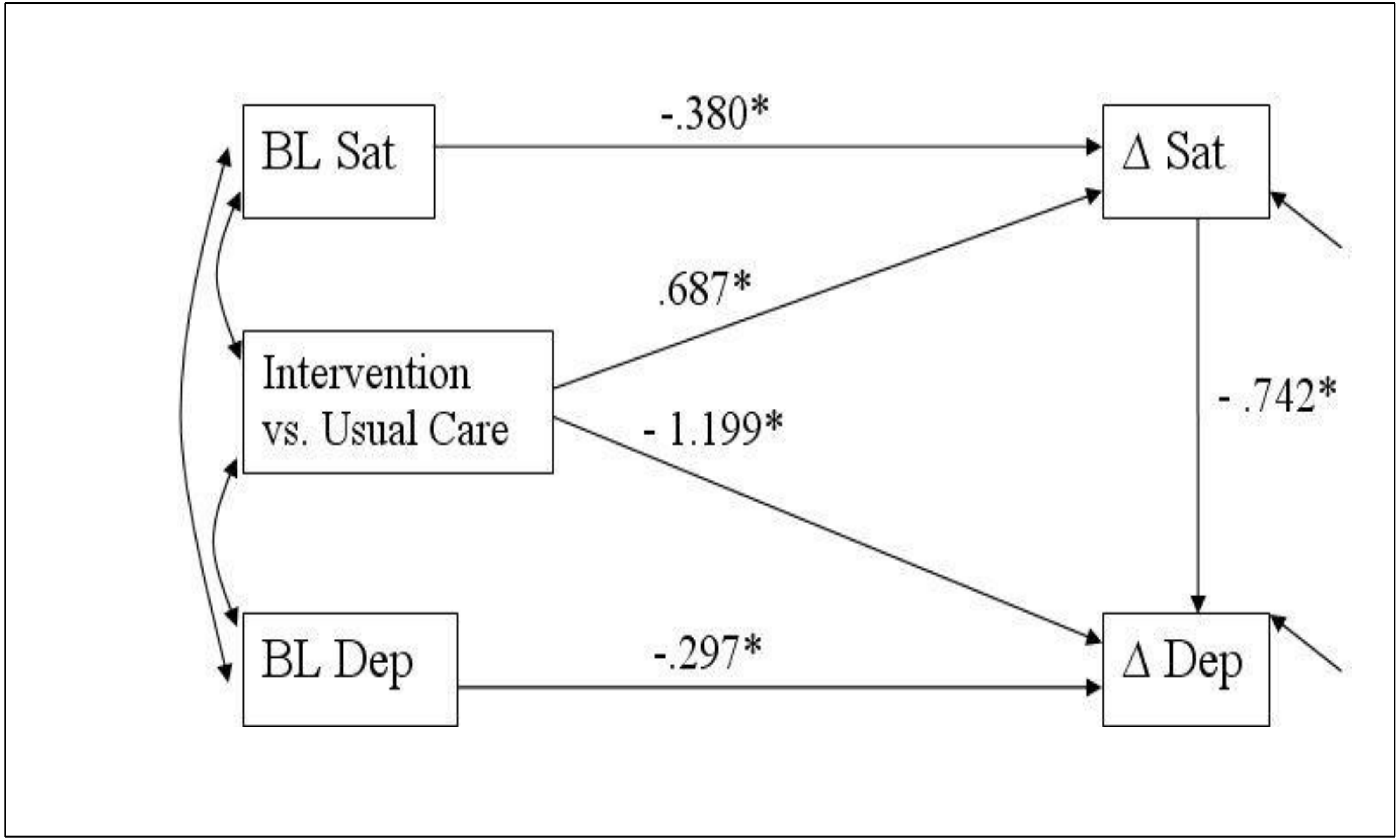
Satisfaction with Social Support in the NYU Caregiver Intervention Study

Likert-type ratings were obtained on how satisfied spouse caregivers were with their social support networks (1 = very dissatisfied, 6 = very satisfied).

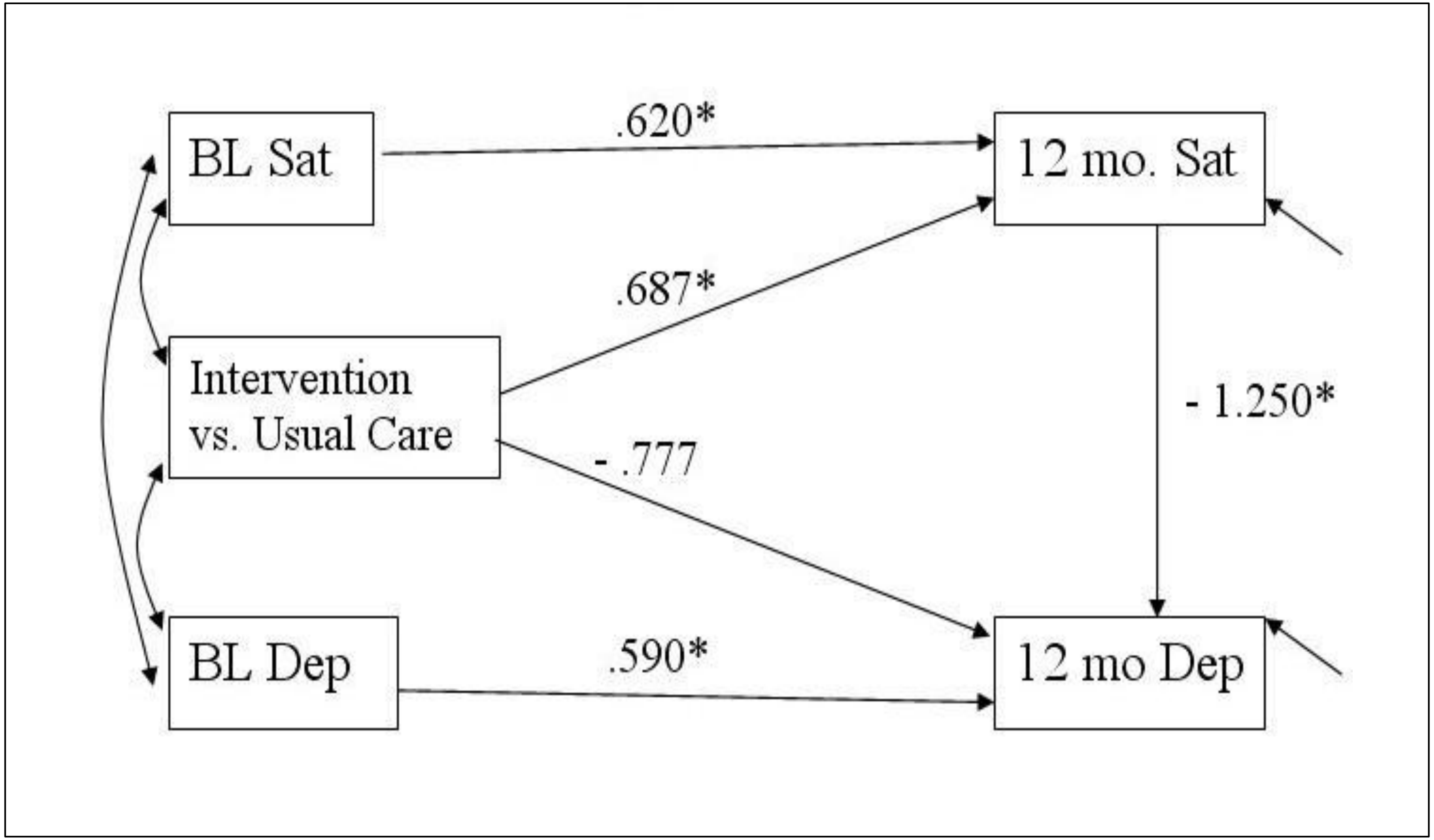
– "In ***general***, how satisfied are you with your social network?"

– "How satisfied are you with the ***assistance*** you get with daily activities (help with chores, patient care)?"

– "How satisfied are you with the ***emotional*** support you get from your social network?"


BL Sat
-.380*
Δ Sat
.687*
Intervention
vs. Usual Care
- 1.199*
- .742*
BL Dep
-.297*
Δ Dep

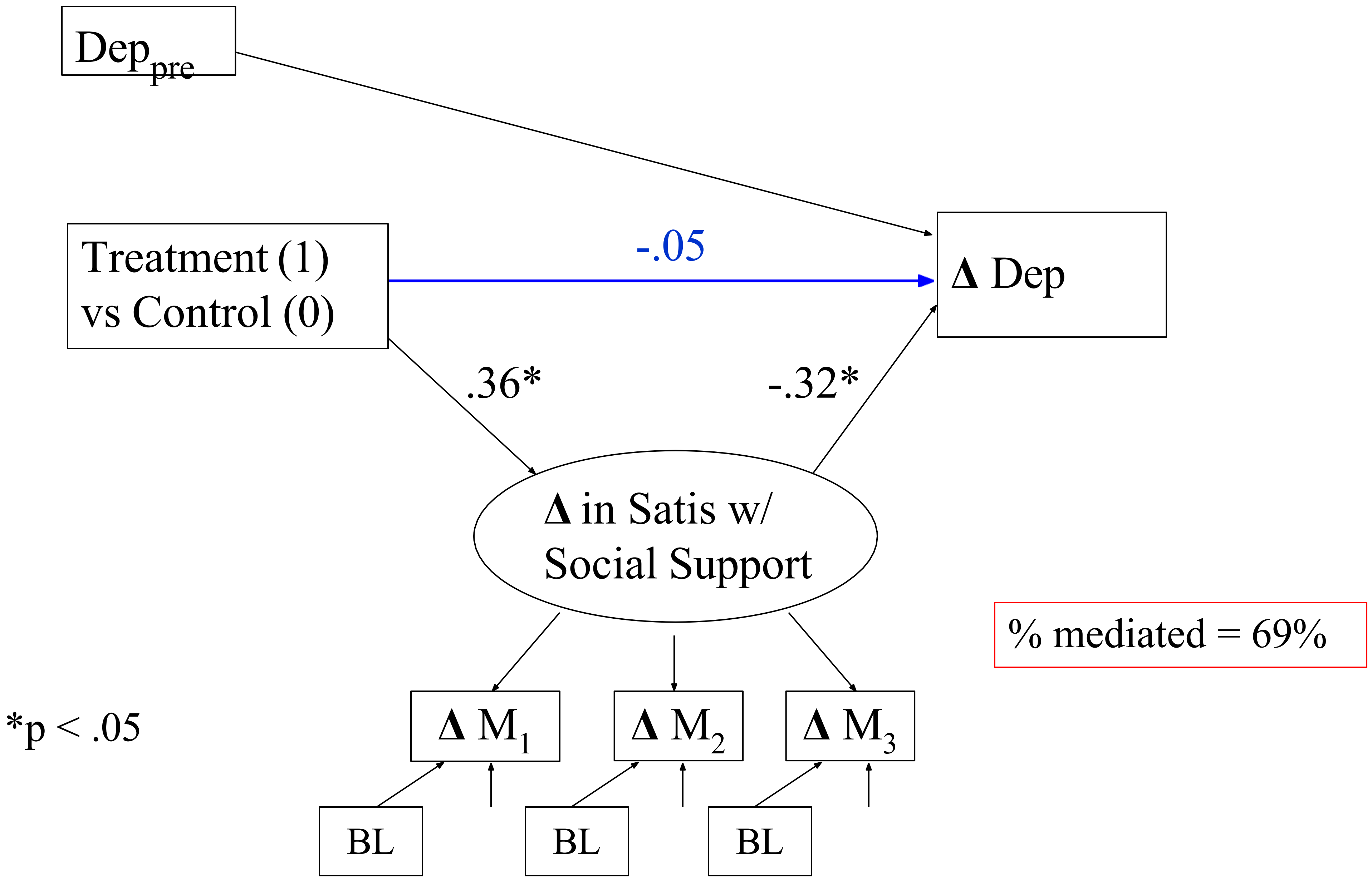
% mediated = 30%


BL Sat
.620*
12 mo. Sat
.687*
Intervention
vs. Usual Care
- 1.250*
- .777
BL Dep
.590*
12 mo Dep

% mediated = 50%

The Case for Latent Mediating Variables

- A key assumption of most multivariate causal models, including ***mediation models***, is that all predictors (including the mediators) are ***measured without error.***
- Monte Carlo simulation studies have shown that even mild ***unreliability*** in the mediating variable can introduce ***serious estimation biases*** (e. g., Hoyle & Kenny, 1999). These biases typically inflate the type II error rate (i.e., reduce power) for the mediated effect (a*b) and inflate the type I error rate (the false positive rate) for the direct effect (c') .
- Latent variables are underlying constructs that are only measured indirectly, usually by their presumed effects on multiple correlated observed indicators. By extracting these common variance components, latent variables are said to be ***measured without error***.
- If you have collected multiple indicators of a mediating process, consider extracting a latent mediating variable.


Dep pre
Treatment (1)
vs Control (0)
-.05
Δ Dep
.36*
-.32*
Δ in Satis w/
Social Support
% mediated = 69%
*p < .05
Δ M1
Δ M2
Δ M3
BL
BL
BL

***Typical Mediation Question*:** **What percentage of an intervention's effect on a primary outcome variable can be explained by that intervention's effect on a mediating mechanism?**

***Answer*:** **Using the exact same data from the NYU caregiver intervention study, we found that the percentage of the intervention's impact on depressive symptoms that was mediated by its effect on satisfaction with social support ranged from *30% to 69%*, depending entirely on the specific analytic method used.**

Improving caregiver well-being delays nursing home placement of patients with Alzheimer disease

Mary S. Mittelman, DrPH; William E. Haley, PhD; Olivio J. Clay, MA; and David L. Roth, PhD

Abstract—*Objective:* To determine the effectiveness of a counseling and support intervention for spouse caregivers in delaying time to nursing home placement of patients with Alzheimer disease (AD), and identify the mechanisms through which the intervention accomplished this goal. *Methods:* We conducted a randomized controlled trial of an enhanced counseling and support intervention compared to usual care. Participants were a referred volunteer sample of 406 spouse caregivers of community-dwelling patients who had enrolled in the study over a 9.5-year period. The intervention consisted of six sessions of individual and family counseling, support group participation, and continuous availability of ad hoc telephone counseling. Structured questionnaires were administered at baseline and at regular follow-up intervals, every 4 months for the first year and every 6 months thereafter. Cox proportional hazard models were used to test the effects of the intervention on the time to nursing home placement for the patients after controlling for multiple time-invariant and time-dependent predictors of placement. *Results:* Patients whose spouses received the intervention experienced a 28.3% reduction in the rate of nursing home placement compared with usual care controls (hazard ratio = 0.717 after covariate adjustment, p = 0.025). The difference in model-predicted median time to placement was 557 days. Improvements in caregivers' satisfaction with social support, response to patient behavior problems, and symptoms of depression collectively accounted for 61.2% of the intervention's beneficial impact on placement. *Conclusion:* Greater access to effective programs of counseling and support could yield considerable benefits for caregivers, patients with Alzheimer disease, and society.

NEUROLOGY 2006;67:1592–1599

eMBRACE
Establishing Mechanisms of Benefit to Reinforce the Alzheimer's Care Experience

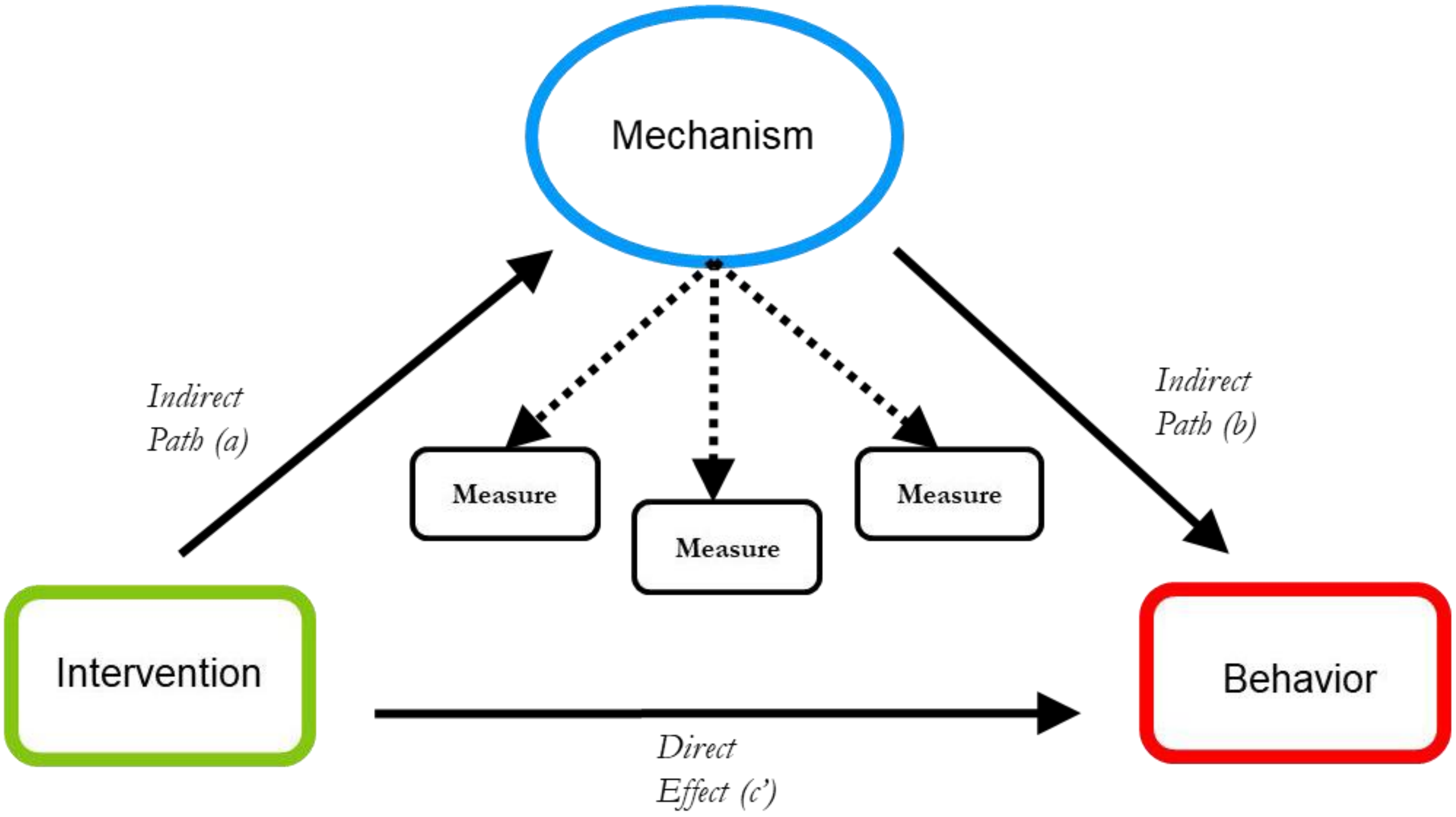
EMORY ROYBAL CENTER FOR DEMENTIA CAREGIVING MASTERY

A guide for conducting rigorous mechanistic research with behavioral interventions: *Introducing the Checklist for Investigating Mechanisms in Behavior-change Research (CLIMBR)*

Jeffrey Birk, PhD
May 22, 2025

COLUMBIA | COLUMBIA UNIVERSITY IRVING MEDICAL CENTER

SOBC Science Of Behavior Change


Mechanism
Indirect Path (a)
Measure
Measure
Measure
Indirect Path (b)
Intervention
Behavior
Direct Effect (c')

Addressing Some Common Challenges

Must I test mediation using a particular **time-ordered relationship** among variables?

- If possible, the mediator should occur in time between the predictor and outcome.
  - Consider the relevant timescales for your research design in terms of expected effects as well as practical considerations.
- You might consider measuring *__changes__* in M and *__changes__* in Y**.**
  - However, a well powered randomized controlled trial does *not* require measurement of M or Y at baseline.
- Entirely cross-sectional research is relatively easy to conduct. However, it may have less utility than a thoughtfully sequenced research design in which the progression of X□M□Y is evaluated over time.

Outline

1. Review the Basic Concepts and Terminology of Statistical Mediation Modeling
2. Applying Mediation Models to Multi-wave (e.g., pre/baseline, post 1, post 2) randomized trials of interventions.
3. **Specific considerations**
   - **A. Analyzing change scores (post – pre) vs. post scores only.**
   - **B. Using baseline values as covariates, when available.**
   - **C. Testing the statistical significance of the mediated effect.**
   - **D. Examining mediation even when the intervention did not have a significant effect on the primary outcome.**
   - **E. The biasing effects of unreliability in mediator measurement. Observed vs. Latent mediators.**
   - **F. Modifications for when the primary outcome is binary, or a time-to-event outcome.**
4. Examples of Applications

Summary of Key Points

3. Specific considerations

   A. Analyzing ***change scores (post – pre)*** vs. post scores only.

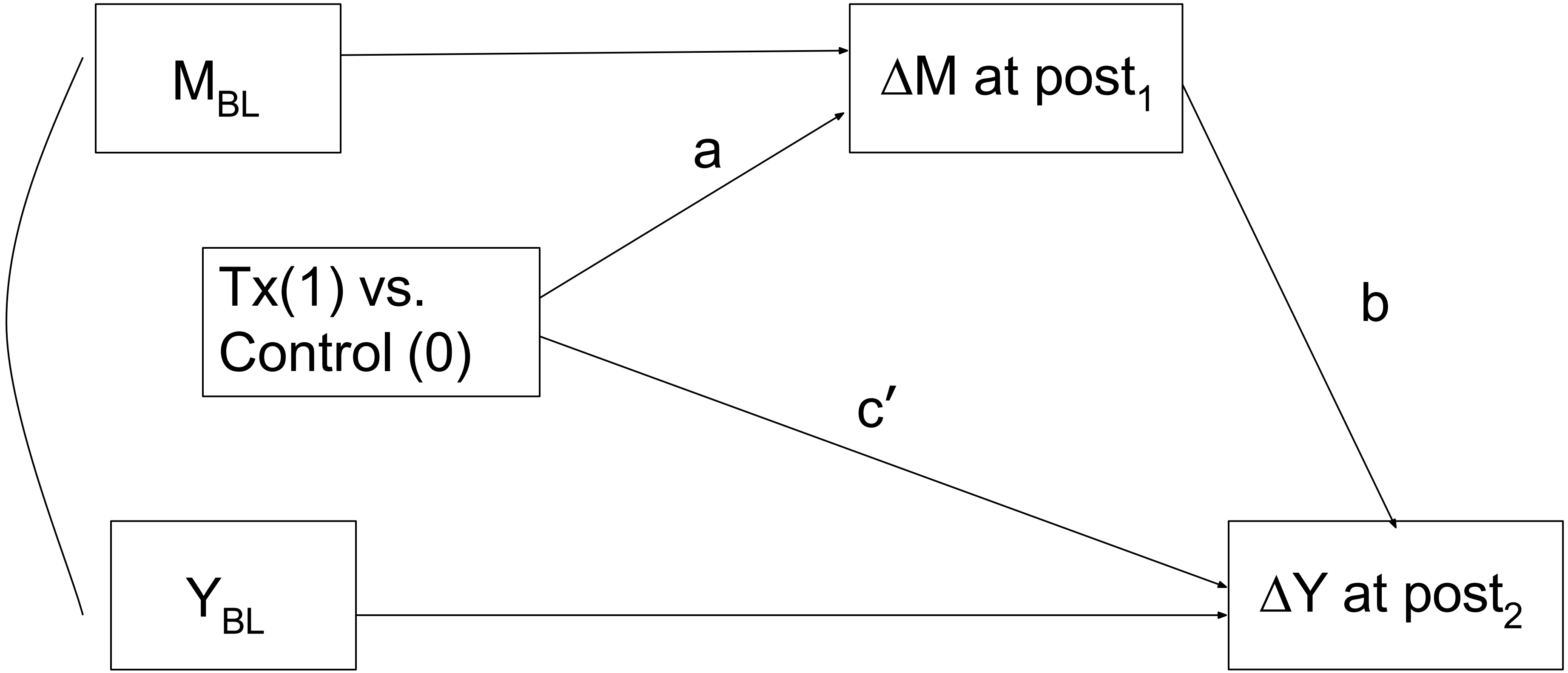
   B. Using ***baseline values as covariates***, when available.

   C. Testing the statistical significance of the mediated effect. ***A bias-corrected bootstrap method usually provides the most power.***

   D. Examining mediation even when the intervention did not have a significant effect on the primary outcome. ***This is fine. The mediated effect may still be statistically significant and interpretable.***

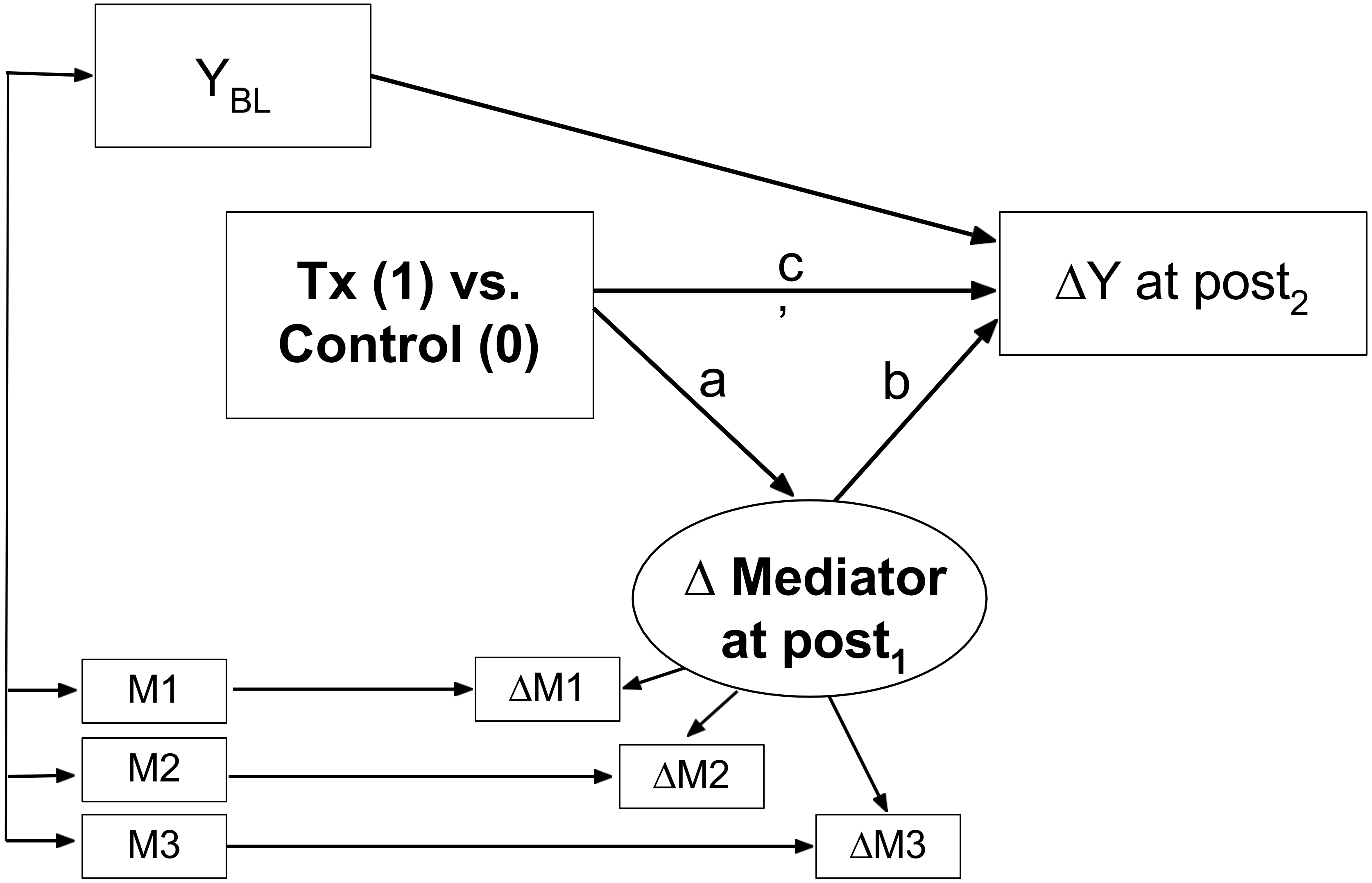
   E. The biasing effects of unreliability in mediator measurement. Observed vs. Latent mediators. ***Unreliability massively undercuts power. Extracting latent mediators, if feasible, is one way to address this. If sticking with observed variables, use measures that have strong psychometric properties.***

Mediation in a 3-wave (Pre, Post$_1$, Post$_2$) RCT Design



where Δ = post – pre change scores.

Mediation in a 3-wave (Pre, $Post_1$, $Post_2$) RCT Design with a Latent Mediator



Methodological References

- Fritz, M. S., & MacKinnon, D. P. (2007). Required sample size to detect the mediated effect. *Psychological Science, 18,* 233-239.
- Hoyle, R. H., & Kenny, D. A. (1999). Sample size, reliability, and tests of statistical mediation. In R. H. Hoyle (Ed.), *Statistical strategies for small sample research*. Thousand Oaks, CA: Sage Publications.
- MacKinnon, D. P., Lockwood, C. M., & Williams, J. (2004). Confidence limits for the indirect effect: Distribution of the product and resampling methods. *Multivariate Behavioral Research, 39*(1), 99-128.
- Muthén, L. K., & Muthén, B. O. (2017). *Mplus: Statistical analysis with latent variables: User's guide*. Author.
- Roth, D. L., & MacKinnon, D. P. (2012). Mediation analysis with longitudinal data. In J. T. Newsom, R. N. Jones, & S. M. Hofer (Eds). *Longitudinal data analysis: A practical guide for researchers in aging, health, and social sciences* (pp. 181-216). Routledge.

Application References

- Diehl, M., Tseng, H. Y., Rebok, G. W., Li, K., Nehrkorn-Bailey, A. M., Rodriguez, D., Chen, D., & Roth, D. L. (2025). Testing the purported mechanisms of the AgingPLUS intervention: Effects on physical activity outcomes. *Psychology and Aging, 40,* 355-370.
- Gitlin, L. N., Roth, D. L., & Huang, J. (2014). Mediators of the impact of a home-based intervention (beat the blues) on depressive symptoms among older African Americans. *Psychology and Aging, 29,* 601-611.
- Gitlin, L. N., Szanton, S. L., Huang, J., & Roth, D. L. (2014). Factors mediating the effects of a depression intervention on functional disability in older African Americans. *Journal of the American Geriatrics Society, 62,* 2280-2287.
- Mittelman, M. S., Haley, W. E., Clay, O. J., & Roth, D. L. (2006). Improving caregiver well-being delays nursing home placement of patients with Alzheimer's disease. *Neurology, 67*, 1592-1599.
- Roth, D. L., Mittelman, M. S., Clay, O. J., Madan, A., & Haley, W. E. (2005). Changes in social support as mediators of the impact of a psychosocial intervention for spouse caregivers of persons with Alzheimer's disease. *Psychology and Aging, 20,* 634-644.

Q & A

Thank you!

Check out our **website**
https://embraceroybal.wisc.edu
or
Contact:
EMBRACE@umn.edu

Establishing Mechanisms of Benefit to Reinforce the Alzheimer's Care Experience

Funding/Disclaimer

Information reported in this presentation was supported by the National Institute On Aging of the National Institutes of Health under Award Number P30AG086642. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.