Partnering with Child Welfare Agencies to Address Contamination Bias and Enhance Scientific Rigor

Chad Shenk, PhD, ABPP

Department of Psychology, University of Rochester
Mt. Hope Family Center, University of Rochester

RESEARCH ARTICLE SUMMARY

PSYCHOLOGY

Estimating the reproducibility of psychological science

Open Science Collaboration*

Is Psychology Suffering From a Replication Crisis?

What Does "Failure to Replicate" Really Mean?

Scott E. Maxwell — *University of Notre Dame*
Michael Y. Lau — *Teachers College, Columbia University*
George S. Howard — *University of Notre Dame*

SCIENCE

Psychology's Replication Crisis Can't Be Wished Away

It has a real and heartbreaking cost.

ED YONG MAR 4, 2016

RESEARCH ARTICLE SUMMARY

PSYCHOLOGY

Estimating the reproducibility of psychological science

Open Science Collaboration*

Is Psychology Suffering From a Replication Crisis?

What Does "Failure to Replicate" Really Mean?

Scott E. Maxwell — *University of Notre Dame*
Michael Y. Lau — *Teachers College, Columbia University*
George S. Howard — *University of Notre Dame*

SCIENCE

Psychology's Replication Crisis Can't Be Wished Away

It has a real and heartbreaking cost.

ED YONG MAR 4, 2016

Child Abuse & Neglect 29 (2005) 479–492

Defining maltreatment according to substantiation: Distinction without a difference?☆

Jon M. Hussey[a,*], Jane Marie Marshall[a], Diana J. English[b], Elizabeth Dawes Knight[c], Anna S. Lau[d], Howard Dubowitz[e], Jonathan B. Kotch[a]

Substantiated and unsubstantiated cases of child maltreatment: Do their consequences differ?

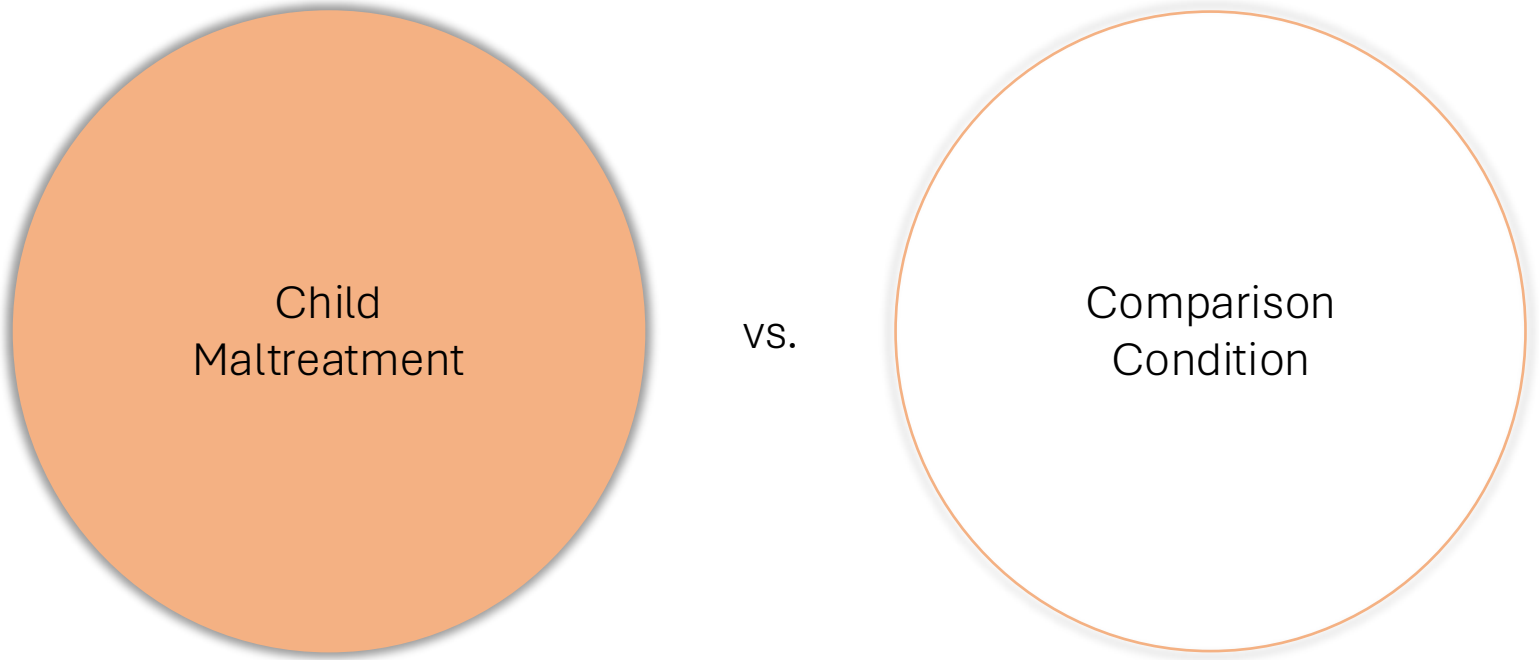
Jeffrey Leiter, Kristen A. Myers, and Matthew T. Zingraff

Child Abuse & Neglect 87 (2019) 112–119

The effect of substantiated and unsubstantiated investigations of child maltreatment and subsequent adolescent health

Kari C. Kugler[a,b], Kate Guastaferro[a], Chad E. Shenk[c], Sarah J. Beal[d], Kathleen M. Zadzora[c], Jennie G. Noll[c,*]

Contamination

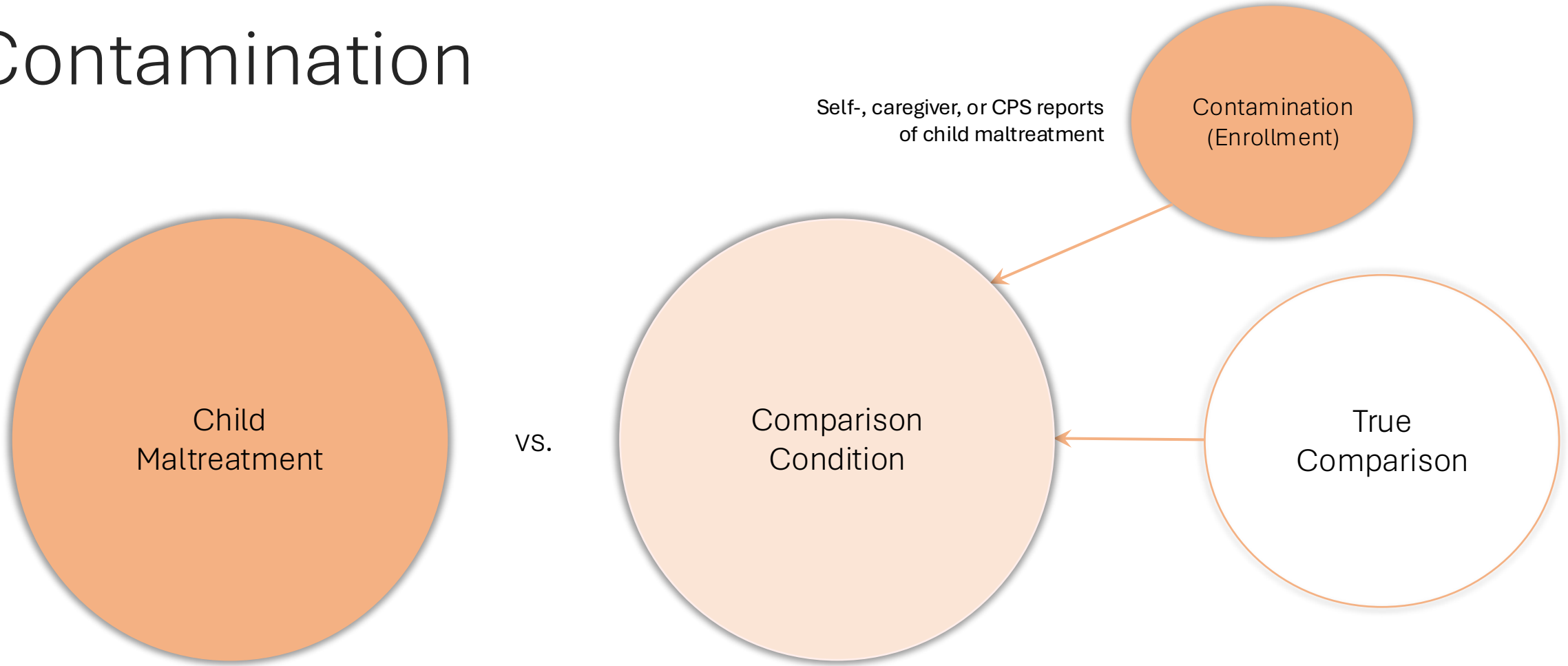


Shenk et al. (2024), *Child Maltreatment*

Contamination

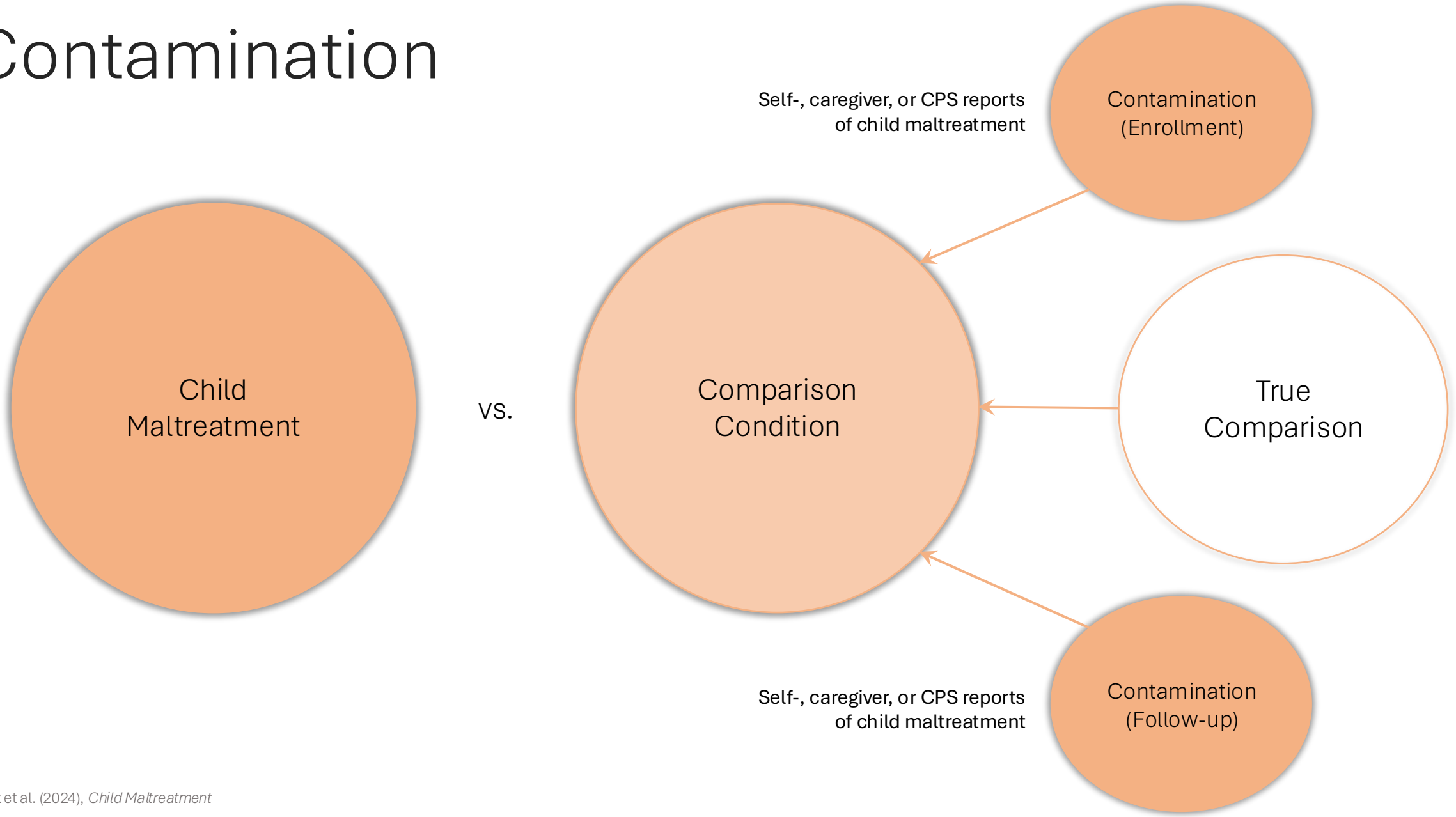
Shenk et al. (2024), *Child Maltreatment*

Contamination



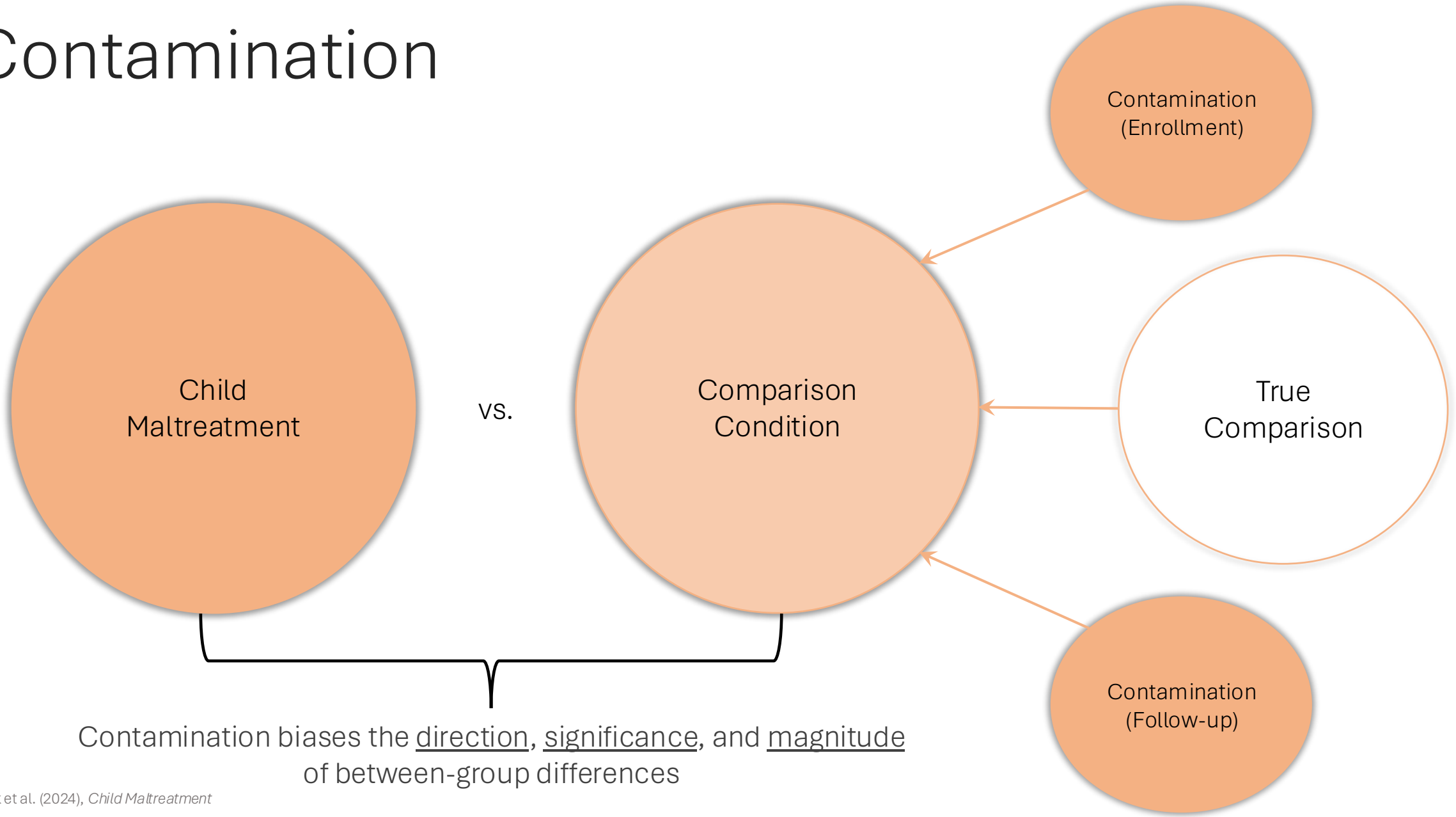
Shenk et al. (2024), *Child Maltreatment*

Contamination



Shenk et al. (2024), *Child Maltreatment*

Contamination



Contamination biases the direction, significance, and magnitude of between-group differences

Shenk et al. (2024), *Child Maltreatment*

Prospectively Ascertained Child Maltreatment and Its Association With *DSM-IV* Mental Disorders in Young Adults

Kate M. Scott, PhD, MA Appl(Clin Psych); Don R. Smith, MA; Pete M. Ellis, PhD, BMBCh

Table 6. *DSM-IV* Disorder Groups Among Young Adults With Child Protection Agency History Compared With Those Without[a]

DSM-IV Mental Disorder Group	Child Protection Agency Group		Comparison Group Including Retrospectively Reported Childhood Maltreatment[b]		Comparison Group			
					Including Retrospectively Reported Childhood Maltreatment[b]		Excluding Retrospectively Reported Childhood Maltreatment[b]	
	No. (%)	SE	No. (%)	SE	Unadjusted OR (95% CI)	Adjusted OR (95% CI)[c]	Unadjusted OR (95% CI)	Adjusted OR (95% CI)[c]
12-Month Disorders								
Any mood disorder	44 (20.94)	3.80	227 (11.90)	1.10	1.96 (1.19-3.23)[d]	1.86 (1.12-3.08)[d]	2.47 (1.47-4.13)[d]	2.38 (1.37-4.14)[d]
Any anxiety disorder	80 (35.74)	5.02	345 (17.92)	1.26	2.55 (1.62-4.00)[d]	2.41 (1.47-3.97)[d]	2.96 (1.87-4.69)[d]	2.92 (1.73-4.91)[d]
Any substance use disorder	38 (16.98)	3.13	186 (8.70)	0.88	2.15 (1.32-3.49)[d]	1.71 (1.01-2.88)[d]	2.55 (1.55-4.21)[d]	2.29 (1.31-4.01)[d]
Any disorder	103 (50.54)	5.03	540 (29.02)	1.81	2.55 (1.62-4.00)[d]	2.32 (1.39-3.85)[d]	2.96 (1.87-4.69)[d]	2.83 (1.68-4.80)[d]
Any 2 disorders	28 (12.13)	3.13	130 (6.11)	0.68	2.12 (1.15-3.92)[d]	1.40 (0.72-2.73)	2.54 (1.35-4.76)[d]	1.55 (0.75-3.21)
Any ≥3 disorders	28 (13.27)	3.03	110 (5.44)	0.68	2.66 (1.49-4.75)[d]	2.67 (1.47-4.87)[d]	3.71 (2.00-6.89)[d]	4.47 (2.29-8.75)[d]
Lifetime Disorders								
Any mood disorder	75 (32.91)	3.95	399 (20.23)	1.25	1.93 (1.31-2.85)[d]	1.80 (1.21-2.68)[d]	2.50 (1.67-3.74)[d]	2.31 (1.52-3.50)[d]
Any anxiety disorder	99 (43.72)	5.22	498 (25.47)	1.46	2.27 (1.47-3.52)[d]	2.04 (1.24-3.33)[d]	2.84 (1.82-2.44)[d]	2.68 (1.61-4.46)[d]
Any substance use disorder	93 (39.68)	4.32	389 (18.15)	1.12	2.97 (2.02-4.37)[d]	2.38 (1.55-3.65)[d]	3.73 (2.51-5.55)[d]	3.11 (1.97-4.91)[d]
Any disorder	136 (64.66)	5.99	795 (43.83)	2.05	2.34 (1.37-4.01)[d]	2.12 (1.20-3.75)[d]	2.91 (1.69-5.00)[d]	2.80 (1.58-4.97)[d]
Any 2 disorders	36 (17.68)	4.04	214 (10.43)	0.88	1.84 (1.05-3.25)[d]	1.30 (0.67-2.53)	2.28 (1.28-4.04)[d]	1.54 (0.77-3.08)
Any ≥3 disorders	67 (28.77)	4.13	247 (11.35)	0.94	3.16 (2.03-4.90)[d]	2.86 (1.79-4.56)[d]	4.19 (2.64-6.66)[d]	3.80 (2.29-6.33)[d]

Contamination Prevalence:

15.1%

Increase in Effect Magnitude:

Past Year Disorder - 22%
Lifetime Disorder - 32%

Scott et al. (2010), *Archives of General Psychiatry*

Contamination in the Prospective Study of Child Maltreatment and Female Adolescent Health

Chad E. Shenk,[1,2] PhD, Jennie G. Noll,[1] PhD, James L. Peugh,[3,4] PhD, Amanda M. Griffin,[1] MS, and Heather E. Bensman,[3] PsyD

Table III. The Relative Risk of Child Maltreatment on Female Adolescent Health

	No control of contamination-baseline model ($N = 498$)		Contamination controlled-multimethod strategy ($N = 394$)	
Outcome	RR	95% CI	RR	95% CI
Teenage births				
Unadjusted	1.80**	1.17–2.78	2.89**	1.47–5.66
Adjusted[a]	1.66*	1.06–2.61	2.21*	1.06–4.63
Obesity				
Unadjusted	1.16	0.88–1.52	1.51*	1.04–2.21
Adjusted[b]	1.16	0.90–1.50	1.47*	1.03–2.08
Major depression				
Unadjusted	1.59	0.99–2.57	4.04**	1.64–9.97
Adjusted[b]	1.28	0.79–2.08	2.95*	1.22–7.16
Past-month cigarette use				
Unadjusted	2.01***	1.51–2.66	2.64***	1.74–3.99
Adjusted[b]	1.36*	1.06–1.74	1.68**	1.21–2.35

*** = p < .001; ** = p < .01; * = p <.05

Shenk et al. (2016), *Journal of Pediatric Psychology*

Contamination Prevalence:

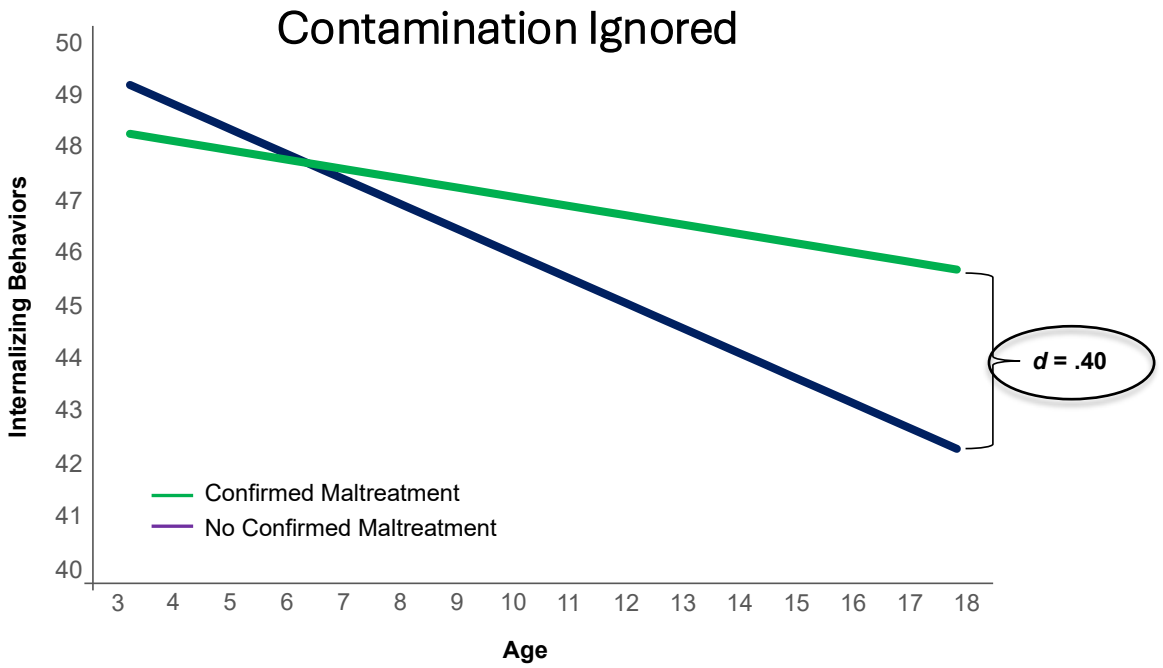
44.8%

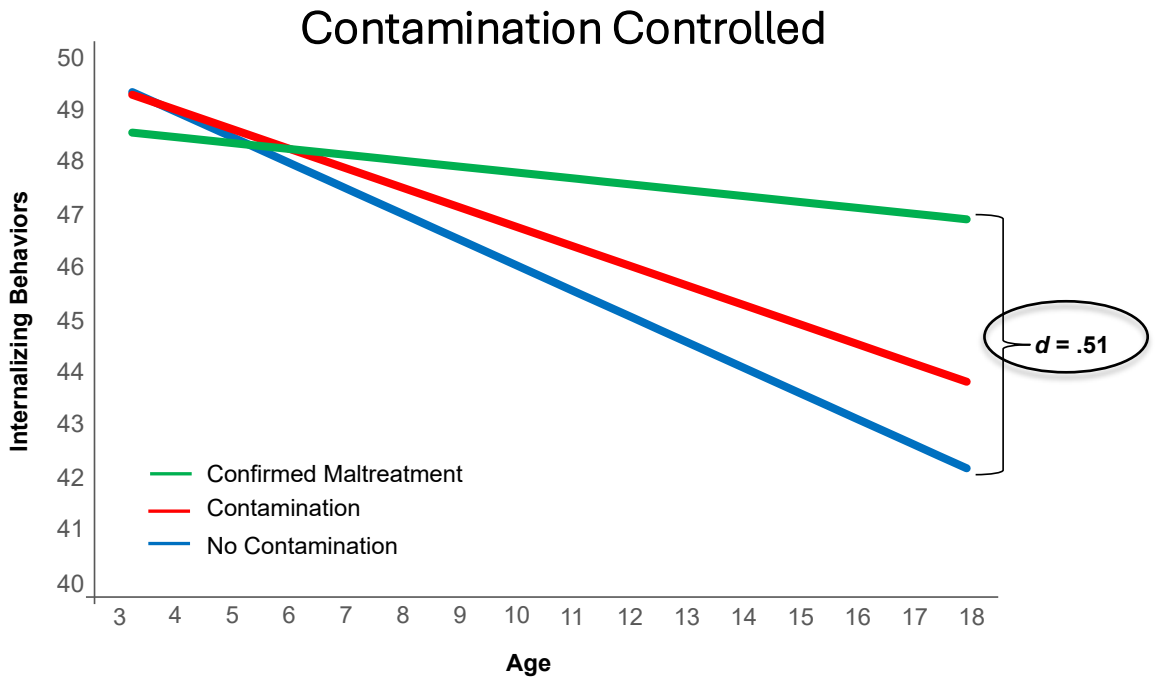
Increase in Effect Size Magnitude:

Teenage Births - 33%
Obesity - 27%
Major Depression - 130%
Past-month Cigarette Use - 24%

Controlling contamination in child maltreatment research: Impact on effect size estimates for child behavior problems measured throughout childhood and adolescence

Chad E. Shenk[1,2], Joseph R. Rausch[3,4], Kenneth A. Shores[5], Elizabeth K. Allen[1] and Anneke E. Olson[1]





Contamination Prevalence:

65.1%

Increase in Effect Size Magnitude:

Internalizing behaviors - 28%
Externalizing behaviors - 53%

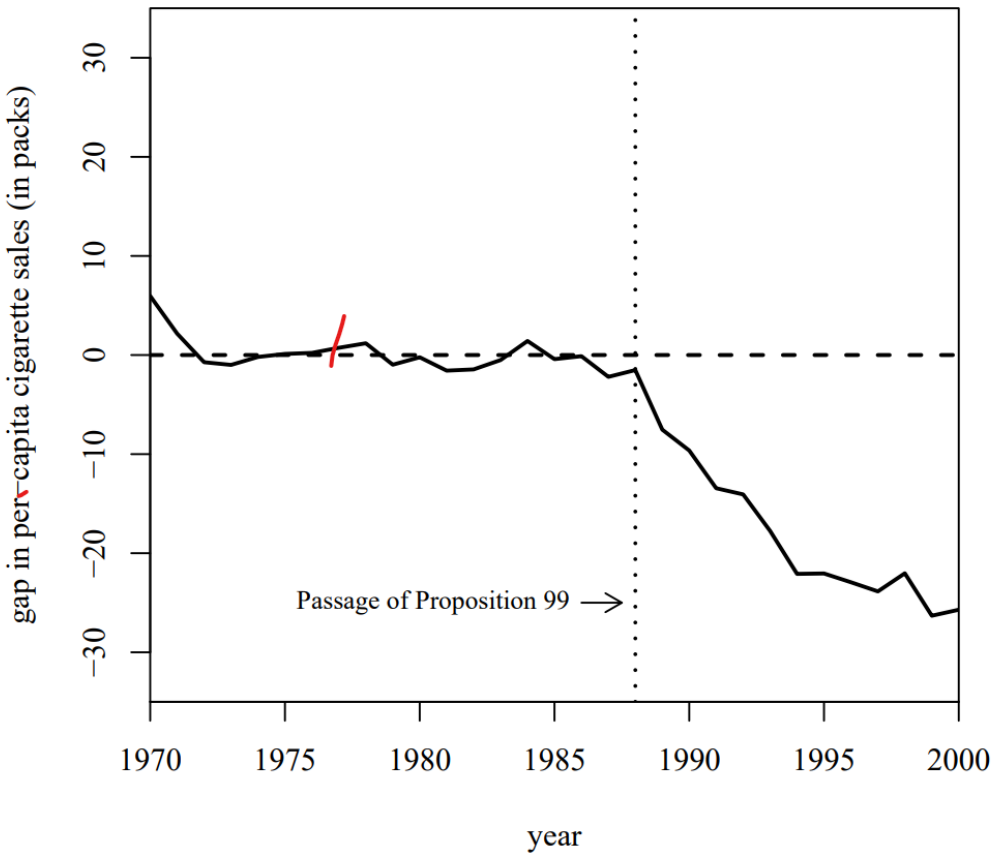
Shenk et al. (2022), *Development and Psychopathology*

Synthetic Control Methods

Average Treatment Effect of the Treated (ATT) is estimated

Estimates are unstandardized mean differences between treated units and synthetic controls

Balances groups on the outcome prior to maltreatment



Abadie et al (2010), *Journal of the American Statistical Association*

Synthetic Control Methods

Longitudinal Studies of Child Abuse and Neglect (LONGSCAN; N=1354)

- Multi-wave, multi-site prospective cohort study in the U.S.
- Confirmed vs. unconfirmed child maltreatment via case record review

Contamination

- Self-report of maltreatment in comparison condition (62%-67%)
- Measured prospectively from age 12, retrospectively at age 18

Trajectories of child behavior problems

- Raw scores for both internalizing and externalizing behaviors
- Measured repeatedly from ages 4 to 18

Shenk, PI (NIH - R03HD104739); Shenk, PI (NSF - BCS-2041333)

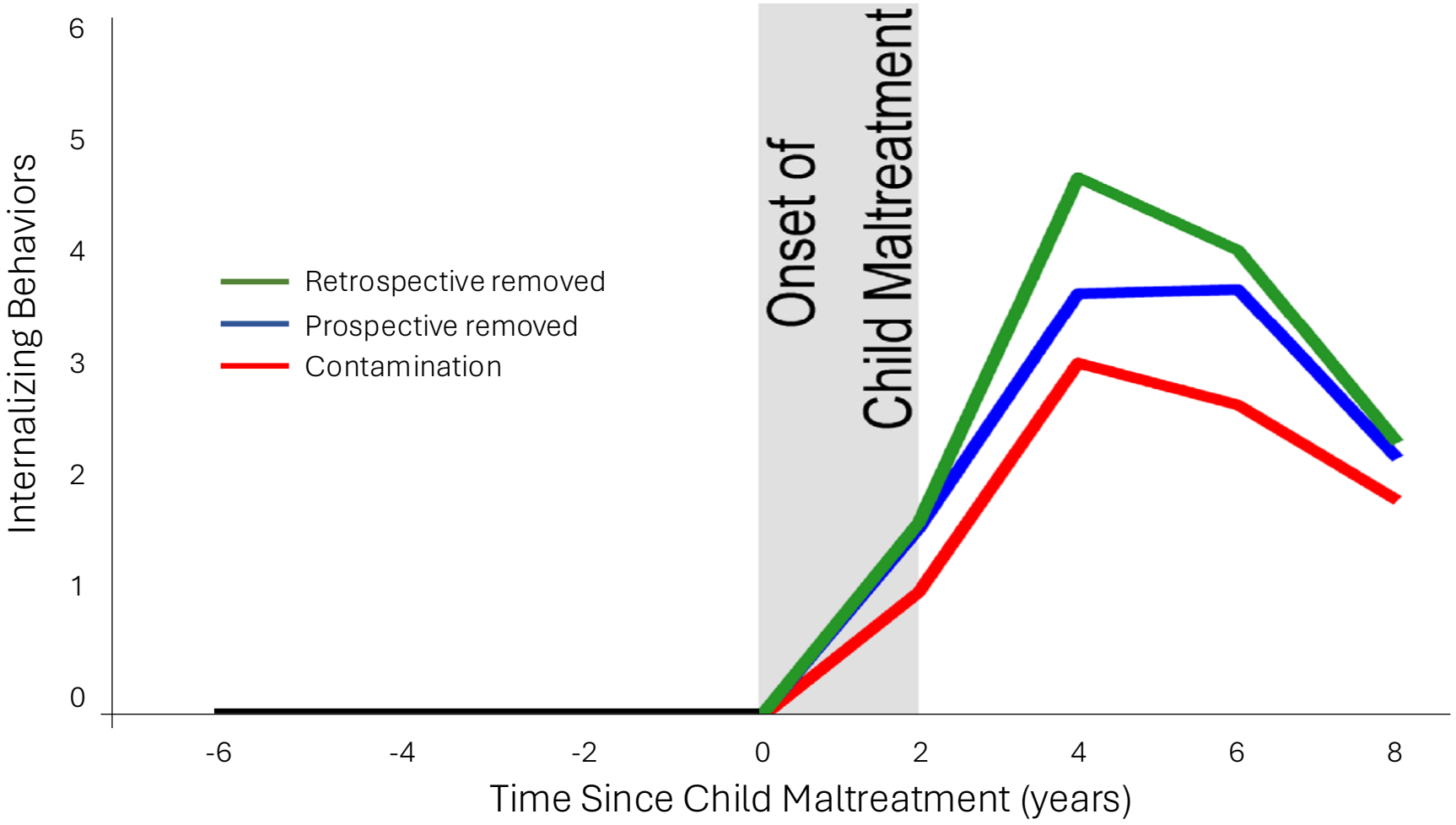
Contamination bias in the estimation of child maltreatment causal effects on adolescent internalizing and externalizing behavior problems

John M. Felt,[1] Ulziimaa Chimed-Ochir,[2] Kenneth A. Shores,[3] Anneke E. Olson,[2] Yanling Li,[2] Zachary F. Fisher,[2] Nilam Ram,[4,5] and Chad E. Shenk[2,6]

Felt et al. (2024), *Journal of Child Psychology and Psychiatry*

Contamination bias in the estimation of child maltreatment causal effects on adolescent internalizing and externalizing behavior problems

John M. Felt,[1] Ulziimaa Chimed-Ochir,[2] Kenneth A. Shores,[3] Anneke E. Olson,[2] Yanling Li,[2] Zachary F. Fisher,[2] Nilam Ram,[4,5] and Chad E. Shenk[2,6]



Direction of Effect:
Increased risk

Significance of Effect:
Ignored - None

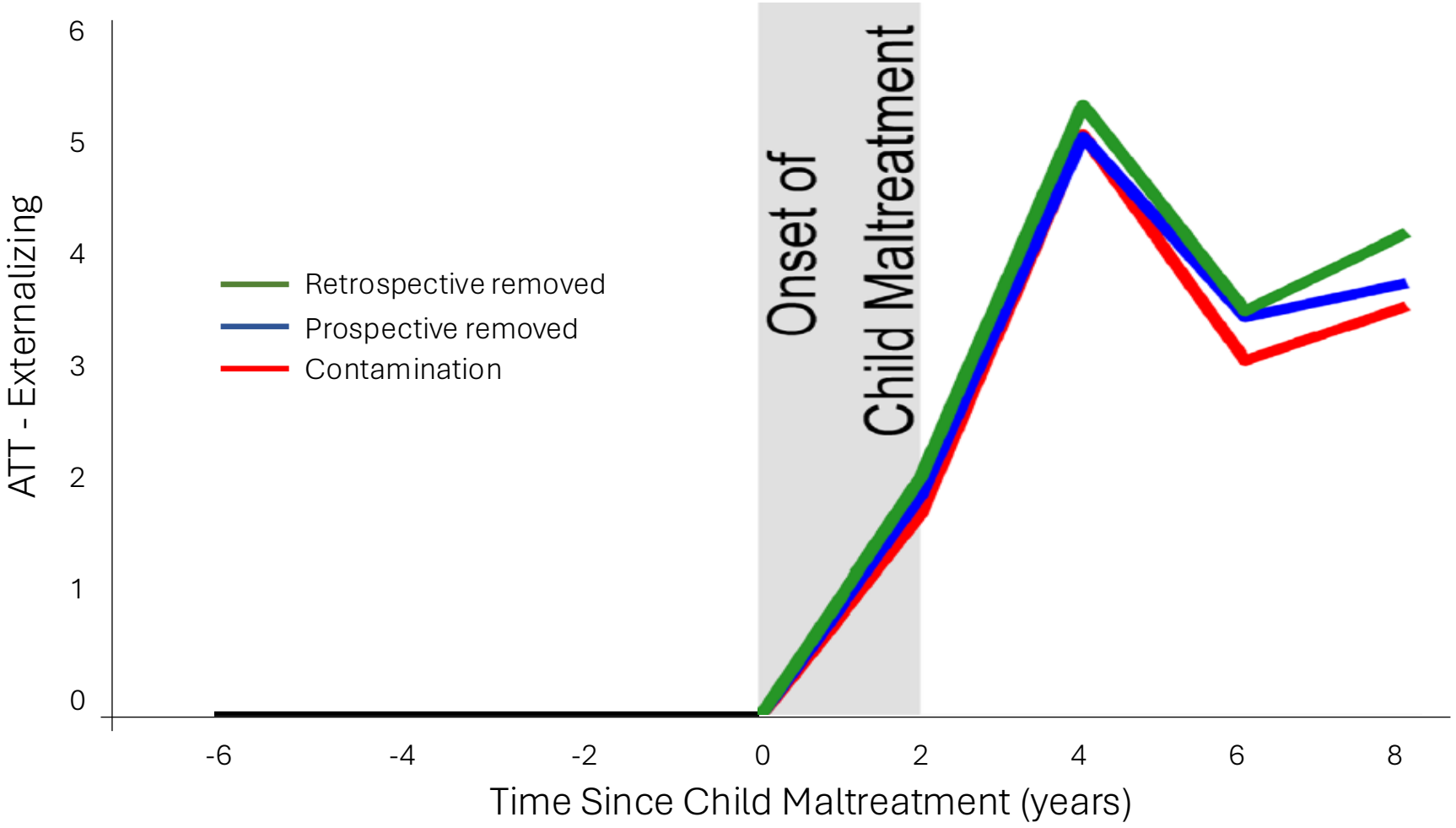
Controlled - **Overall, 2 and 4 years** post-maltreatment

Magnitude of Effect:
Increases of 20%-52%

Felt et al. (2024), *Journal of Child Psychology and Psychiatry*

Contamination bias in the estimation of child maltreatment causal effects on adolescent internalizing and externalizing behavior problems

John M. Felt,[1] Ulziimaa Chimed-Ochir,[2] Kenneth A. Shores,[3] Anneke E. Olson,[2] Yanling Li,[2] Zachary F. Fisher,[2] Nilam Ram,[4,5] and Chad E. Shenk[2,6]



Direction of Effect:
Increased risk

Significance of Effect:
Ignored - 2 years

Controlled - **Overall, 2 years** post-maltreatment

Magnitude of Effect:
Increases of 0%-18%

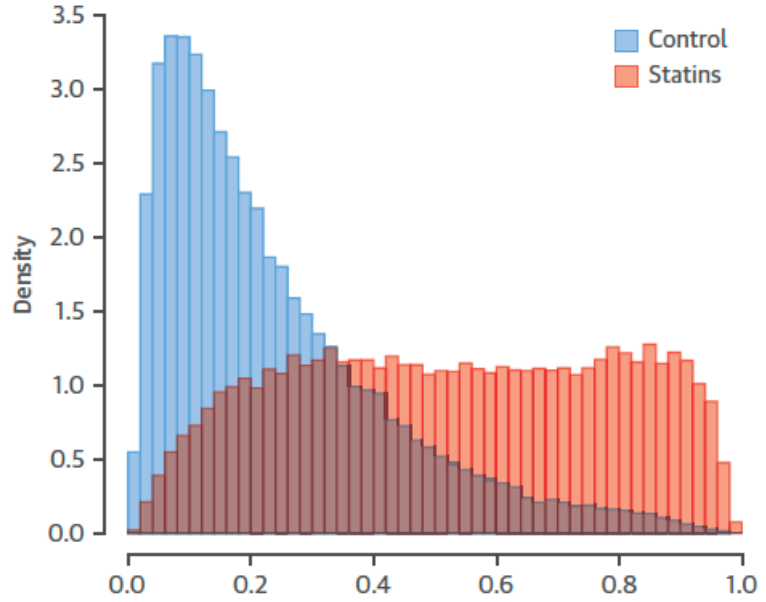
Felt et al. (2024), *Journal of Child Psychology and Psychiatry*

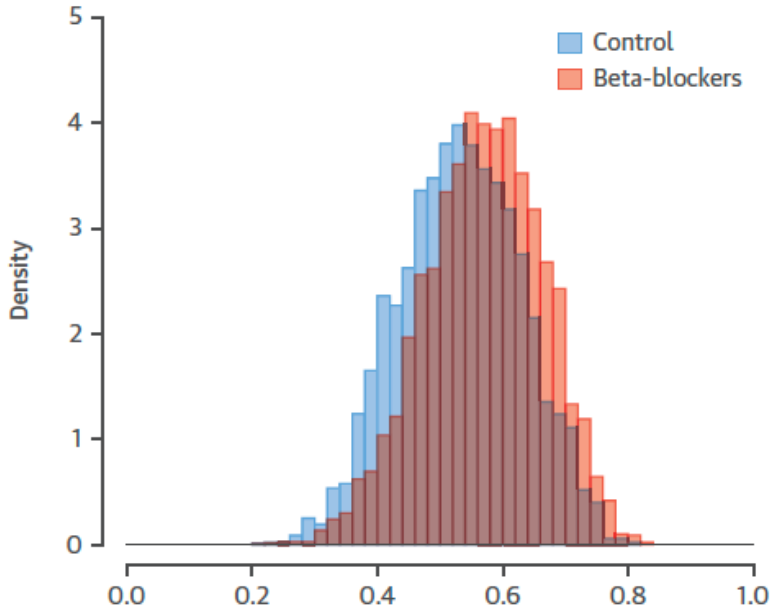
Propensity Score Methods

Average Treatment Effect of the Population (ATE) can be estimated

Estimates are unstandardized mean differences on an outcome between maltreated and control conditions

Balances groups on a vector of covariates prior to treatment





Rosenbaum & Rubin (1983), *Biometrika*; Elze et al (2017), *Journal of the American College of Cardiology*

Propensity Score Methods

National Study on Child and Adolescent Well-being-II (NSCAW-II; N=5872)

- U.S. national probability sample of child welfare population
- Substantiated vs. unsubstantiated child maltreatment

Contamination

- Caregiver-report of maltreatment in comparison condition (96.1%)
- Modeled as a third level of the treatment

Child behavior problems

- *T*-scores for internalizing and externalizing behaviors 3 years post

Shenk, PI (R03HD104739); Shenk, PI (NSF-BCS-2041333)

Propensity Score Methods

Contamination Ignored

Internalizing

Parameter	ATE	95% CI
Intercept	52.00	51.40, 52.70
Substantiated	-0.68	-1.59, 0.23

Externalizing

Parameter	ATE	95% CI
Intercept	54.00	53.30, 54.70
Substantiated	-0.98*	-1.93, -0.04

Substantiated vs. Unsubstantiated?

Olson et al. (under review)

Propensity Score Methods

Contamination Ignored

Internalizing

Parameter	ATE	95% CI
Intercept	52.00	51.40, 52.70
Substantiated	-0.68	-1.59, 0.23

Externalizing

Parameter	ATE	95% CI
Intercept	54.00	53.30, 54.70
Substantiated	-0.98*	-1.93, -0.04

Contamination Controlled

Internalizing

Parameter	ATE	95% CI
Intercept	41.06	38.40, 43.70
Substantiated	10.29***	7.56, 13.08

Externalizing

Parameter	ATE	95% CI
Intercept	44.60	39.40, 49.70
Substantiated	8.46**	3.28, 13.60

*** = p < .001; ** = p < .01; * = p <.05

Olson et al. (under review)

	Prospective studies*	Retrospective studies*
Education and employment		
Low educational achievement	Moderate	Weak
Low skilled employment	Moderate	Lacking
Mental health		
Behaviour problems as child/adolescent	Strong	Strong
Post-traumatic stress disorder	Strong	Strong
Depression	Moderate	Strong
Attempted suicide	Moderate	Strong
Self-injurious behaviour	Weak	Weak
Alcohol problems	Moderate	Strong
Drug misuse/dependence	Weak	Strong
Physical health and sexual behaviour		
Prostitution/sex trading	Moderate	Strong
Teenage pregnancy	Inconsistent	Strong
Promiscuity	No effect	Strong
General adult health	Lacking	Moderate
Chronic pain in adulthood	No effect	Weak
Obesity	Strong	Weak
Health-care use/costs	Lacking	Moderate
Quality of life	Lacking	Lacking
Aggression, violence, criminality		
Criminal behaviour	Strong	Strong

*Refers to ascertainment of maltreatment. The classification indicates consensus about the findings from included studies and are broadly consistent with the following criteria: strong=evidence of a significant effect after adjustment for confounders; moderate=evidence of a significant but small effect, or of a stronger effect that is reduced after adjustment for confounders or highly likely to be confounded; weak=evidence of an effect based on methodologically problematic studies or associations that do not persist after adjustment, but consistently favour a positive effect; inconsistent=effect qualitatively different across studies (ie, positive and negative or no associations); lacking=no studies addressing this question.

Table 2: **Summary of review findings on consequences of child maltreatment—evidence for an association in prospective and retrospective studies**

Gilbert et al. (2009), *Lancet*

Failure to Replicate

Bias in the direction and significance of estimates across studies due to varying degrees of contamination

	Prospective studies*	Retrospective studies*
Education and employment		
Low educational achievement	Moderate	Weak
Low skilled employment	Moderate	Lacking
Mental health		
Behaviour problems as child/adolescent	Strong	Strong
Post-traumatic stress disorder	Strong	Strong
Depression	Moderate	Strong
Attempted suicide	Moderate	Strong
Self-injurious behaviour	Weak	Weak
Alcohol problems	Moderate	Strong
Drug misuse/dependence	Weak	Strong
Physical health and sexual behaviour		
Prostitution/sex trading	Moderate	Strong
Teenage pregnancy	Inconsistent	Strong
Promiscuity	No effect	Strong
General adult health	Lacking	Moderate
Chronic pain in adulthood	No effect	Weak
Obesity	Strong	Weak
Health-care use/costs	Lacking	Moderate
Quality of life	Lacking	Lacking
Aggression, violence, criminality		
Criminal behaviour	Strong	Strong

*Refers to ascertainment of maltreatment. The classification indicates consensus about the findings from included studies and are broadly consistent with the following criteria: strong=evidence of a significant effect after adjustment for confounders; moderate=evidence of a significant but small effect, or of a stronger effect that is reduced after adjustment for confounders or highly likely to be confounded; weak=evidence of an effect based on methodologically problematic studies or associations that do not persist after adjustment, but consistently favour a positive effect; inconsistent=effect qualitatively different across studies (ie, positive and negative or no associations); lacking=no studies addressing this question.

Table 2: **Summary of review findings on consequences of child maltreatment—evidence for an association in prospective and retrospective studies**

Gilbert et al. (2009), *Lancet*

Failure to Replicate

Bias in the direction and significance of estimates across studies due to varying degrees of contamination

Even if risks are "significant", effect sizes will vary across studies depending on different levels of contamination and whether contamination is controlled

JAMA Network Open

Original Investigation | Substance Use and Addiction

Association Between Daily Alcohol Intake and Risk of All-Cause Mortality
A Systematic Review and Meta-analyses

Jinhui Zhao, PhD; Tim Stockwell, PhD; Tim Naimi, MD; Sam Churchill, MSc; James Clay, MSc; Adam Sherk, PhD

Table 2. Mean Relative Risk Estimates of All-Cause Mortality Due to Alcohol Consumption Up to 2022 According to 107 Studies With 724 Relative Risk Estimates

Drinking categories	Studies, No./risk estimates, No.	Unadjusted[a]		Partially adjusted[b]		Fully adjusted[c]	
		RR (95% CI)	*P* value	RR (95% CI)	*P* value	RR (95% CI)	*P* value
Reference group = lifetime nondrinker							
Abstainer	107/191	1 [Reference]		1 [Reference]		1 [Reference]	
Any drinker vs abstainer	107/724	1.06 (0.90-1.25)	.42	1.03 (0.89-1.19)	.65	1.11 (0.96-1.28)	.12
Former drinker vs abstainer	28/56	1.22 (1.11-1.33)	<.001	1.17 (1.08-1.26)	<.001	1.26 (1.12-1.42)	.0001
Active drinker vs abstainer, g/d	107/668	0.97 (0.94-1.00)	.02	0.93 (0.90-0.96)	<.001	1.02 (0.93-1.13)	.61
Occasional (<1.30)	24/57	0.92 (0.84-1.01)	.08	0.89 (0.83-0.95)	<.001	0.96 (0.86-1.06)	.41
Low-volume (1.30 to <25)	99/306	0.85 (0.81-0.88)	<.001	0.86 (0.83-0.88)	<.001	0.93 (0.85-1.01)	.08
Medium volume (25 to <45)	80/146	1.02 (0.96-1.08)	.55	0.97 (0.92-1.02)	.21	1.05 (0.96-1.14)	.28
High volume (45 to <65)	52/76	1.07 (0.99-1.16)	.09	1.11 (1.03-1.21)	.009	1.19 (1.07-1.32)	.001
Higher volume (≥65)	45/83	1.35 (1.26-1.46)	<.001	1.24 (1.16-1.32)	<.001	1.35 (1.23-1.47)	.0001

"The reduced RR estimates for occasional or moderate drinkers observed without adjustment may be due to the misclassification of former and occasional drinkers into the reference group…"

"Of 107 studies identified, 86 included former drinkers and/or occasional drinkers in the abstainer reference group…"

Contamination Prevalence:

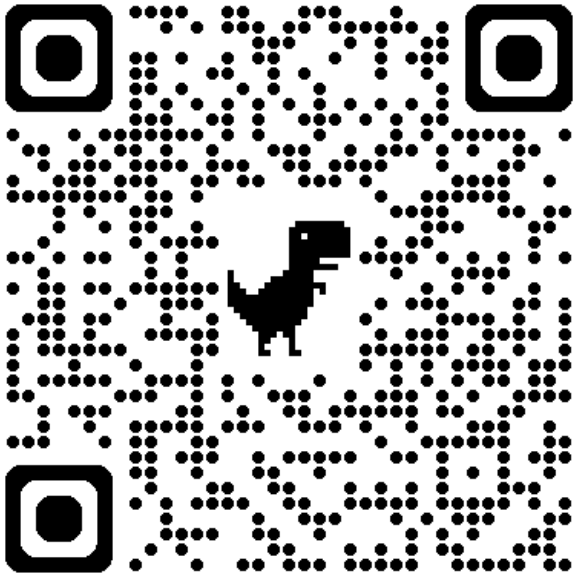
80% across studies

Thank you!

Email: chad.shenk@rochester.edu

Social: @ChadShenkPhD

Web:



Funders:

NIH National Institutes of Health
Turning Discovery Into Health

NSF

American Psychological Association

Collaborators:

Kenneth Shores, PhD

Nilam Ram, PhD

Zachary Fisher, PhD

John Felt, PhD

Anneke Olson, PhD