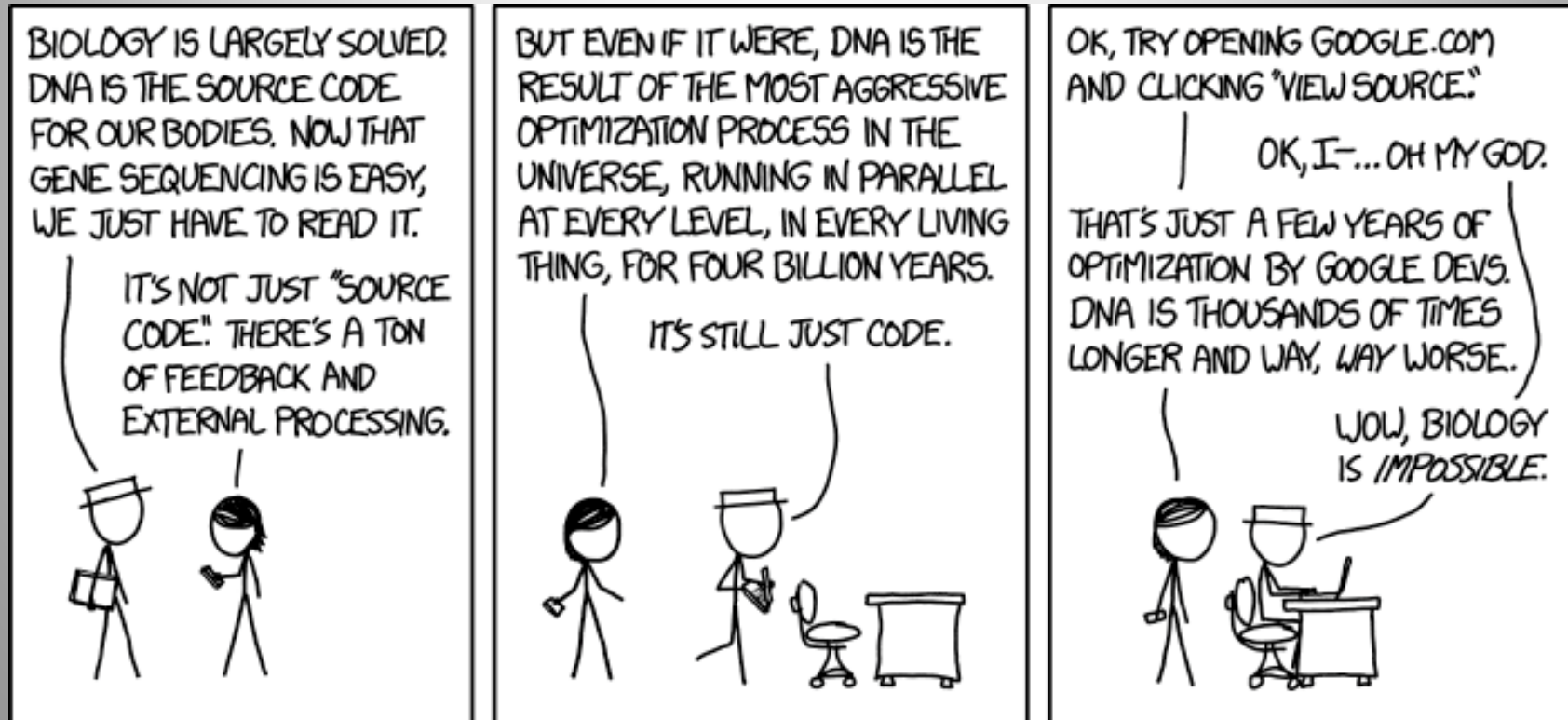


# Causation Models in Twin Studies

Brad Verhulst



# Experimentation: The Gold Standard for Assessing Causality

The Logic of Experimentation:

1. Take individuals drawn at random from a population
  - If you have selected (ascertained) samples some of the logic may not be sound and you may (probably will) have bias parameter estimates.
  - Convenience samples (such as undergrads) are not inherently problematic, but they may impede generalizability
2. Randomly assign them to one of two (or more) groups
  - Everyone had the same opportunity to be in each group, so there should be no meaningful differences between groups
  - This requires groups to be of sufficient size to allow for random variation in the population to be distributed across the groups
3. Expose Group A to Treatment A and Group B to Treatment B
  - Such as Drug vs Placebo, treatment vs control ...
4. If there are any post treatment differences between Group A and Group B, then the difference must be caused by the difference in treatments

# But ...

Many of the questions that we are interested in cannot be manipulated

Research Question:

- Do extraverted individuals have higher levels of substance use initiation?
  - Difficult to manipulate Extraversion
- Does Childhood Maltreatment increase the risk of depression?
  - Probably shouldn't abuse children to test scientific hypotheses

# Discordant MZ Twin Design

- MZ twins share a high level of genetic and environmental factors
  - Probably better than any propensity score matching algorithm that you could ever imagine
- If something randomly happened to one MZ twin, subsequent differences between twins may be attributable to this random event.

# ***Early Exposure to Anesthesia and Learning Disabilities in a Population-based Birth Cohort***

Robert T. Wilder, M.D., Ph.D.,\* Randall P. Flick, M.D., M.P.H.,† Juraj Sprung, M.D., Ph.D.,‡ Slavica K. Katusic, M.D.,§ William J. Barbaresi, M.D.,|| Christopher Mickelson, M.D.,# Stephen J. Gleich, M.D.,\*\* Darrell R. Schroeder, M.S.,†† Amy L. Weaver, M.S.,†† David O. Warner, M.D.‡

**Table 6. Effects of Anesthetic Exposures before Age 4 yr on Risk for Development of Learning Disabilities**

	Unadjusted			Adjusted*		
	Hazard Ratio	95% CI	P Value	Hazard Ratio	95% CI	P Value
Number of exposures			< 0.001			< 0.001
0, n = 4,764	Reference			Reference		
1, n = 449	1.05	0.84–1.32		1.00	0.79–1.27	
2, n = 100	1.78	1.22–2.59		1.59	1.06–2.37	
3 or more, n = 44	2.50	1.55–4.04		2.60	1.60–4.24	

**Strong Evidence for an association between Early Anesthesia Exposure and Learning Disabilities**

**Does Early Anesthesia Exposure cause Learning Disabilities?**

# Anesthesia and Cognitive Performance in Children: No Evidence for a Causal Relationship

Meike Bartels,<sup>1\*</sup> Robert R. Althoff,<sup>2\*</sup> and Dorret I. Boomsma<sup>1</sup>

EA exposure	DIS_NE	DIS_E
Under 3: males*	536.76	536.40
Ever: males**	538.54	538.83
Under 3: females***	534.43	534.36
Ever: females	536.26	535.69
CP	DIS_NE	DIS_E
Under 3: males	2.98	2.75
Ever: males	2.90	2.73
Under 3: females*	4.26	3.44
Ever: females*	2.86	2.67

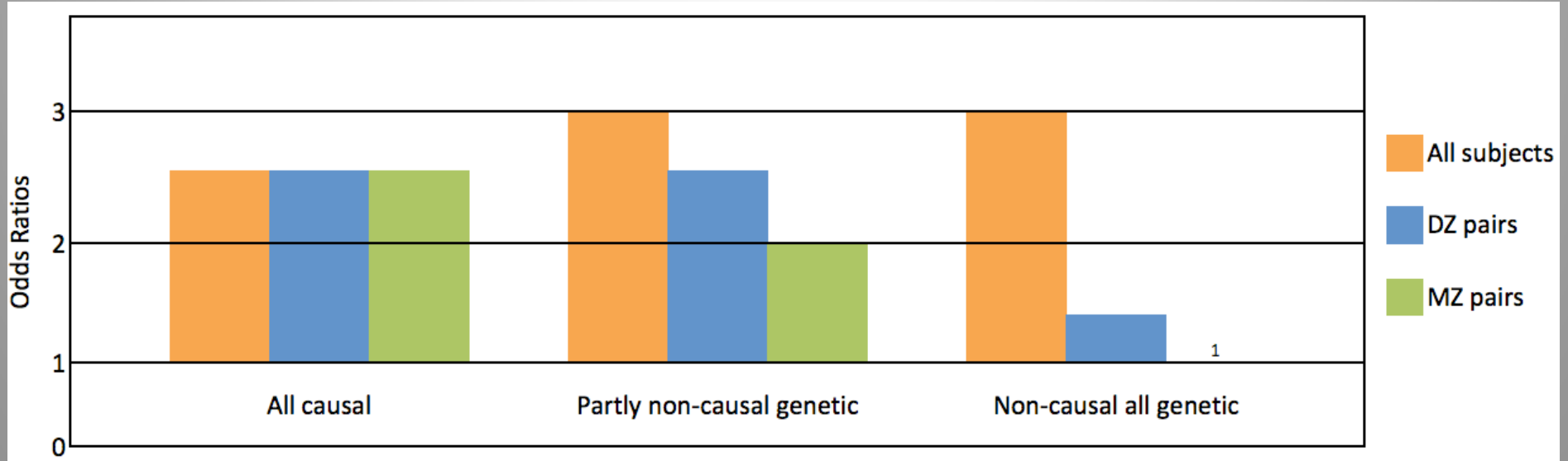
There are no differences between MZ twins based on Early Exposure to Anesthesia Exposure

Okay, but there are DZ twins too



While DZ twins do not share the same degree of genetic relatedness, they can still tell us something about the casual relationship between two variables

- **If risk factor-outcome association is:**
- Causal
  - Controlling for background & genetic effects makes no difference
  - Estimates (ORs) are same
- Partly due to G factors influencing risk & outcome
  - Association strongest in entire sample (no control),
  - Intermediate for DZs (full E, part G control)
  - Lowest for MZ (full G & E control)
- Entirely genetic
  - MZ ORs approach 1, DZ ORs are midway



Courtesy of Gillespie & Kendler

## **Commentary: Direction of Causation Models**

**Jack Goldberg and Viswanathan Ramakrishnan**

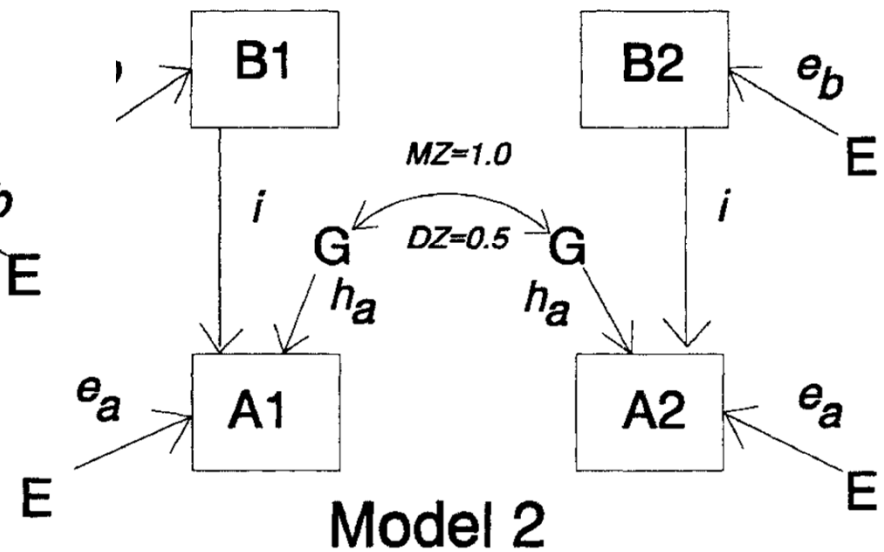
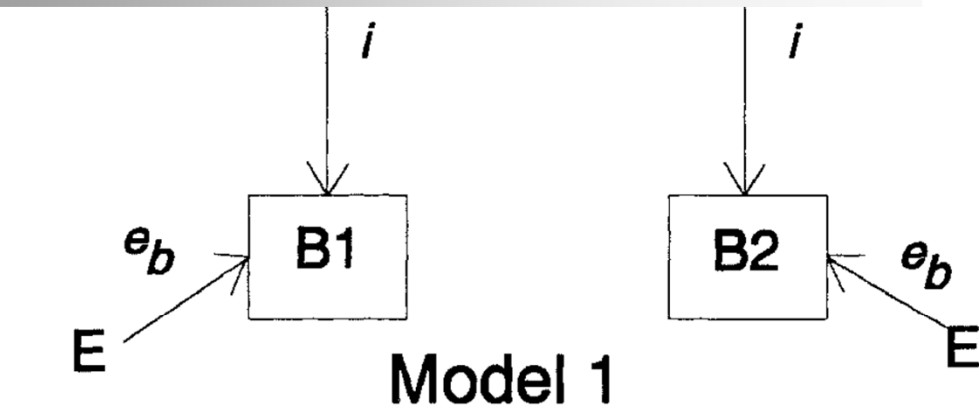
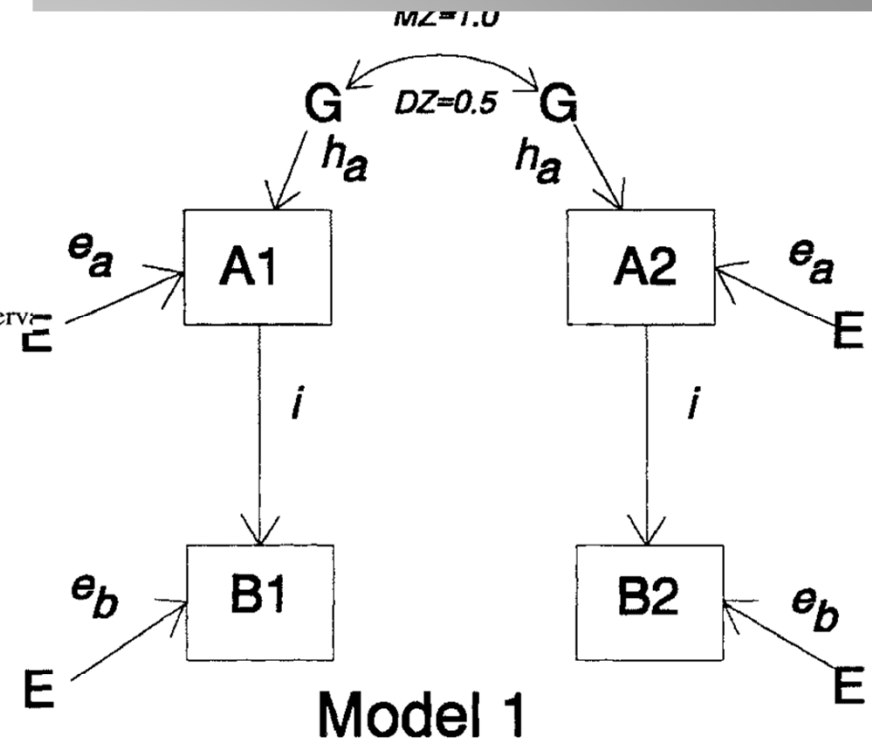
limitations of the method. To many epidemiologists without training in the intricacies of the biometrical genetic analysis of twin data the approach may appear mysterious. These epidemiologists are likely to reject the approach without adequately considering the merits of the method. We feel that this would be a serious error since the direction of causation models are of potentially great value.

known. Failure to correctly specify a measurement model can lead to incorrect tests of hypotheses. Difficulties can also occur when discriminating between a direct causative relationship and a correlation due to common genetic or environmental determinants, but these occur in predictable situations. If these considerations are taken into account in interpretation of results, the true nature of the association between traits can often be correctly identified, or at least included in a subgroup of best fitting models. © 1994 Wiley-Liss, Inc.

**Key words:** twins, causation, smoking, alcohol intake, lung function

**INTRODUCTION**

A well known shortcoming of cross-sectional and even longitudinal observ



# Direction of Causality: A Comment

Gregory Carey

*Institute for Behavioral Genetics and Department of Psychology, University of Colorado, Boulder*

	A1	A2	B1	B2
A1				
A2	$\rho_A$			
B1	$i$	$t$		
B2	$t$	$i$	$\rho_B$	

In Siblings (not twins):

If  $A \rightarrow B$ :  $t = i\rho_A$

If  $B \rightarrow A$ :  $t = i\rho_B$

If  $\rho_A = \rho_B$  then both models ( $A \rightarrow B$  &  $B \rightarrow A$ ) give the same expectations

If there are differences in the modes of transmission (which you know because you have twin data), then you can resolve this issue.

In the case where  $A \rightarrow B$  (A is AE and B is AE) the expected MZ and DZ correlation matrices would be:

	MZ twins				DZ twins			
	A1	A2	B1	B2	A1	A2	B1	B2
A1	1				1			
A2	$h_a^2$	1			$\frac{1}{2}h_a^2$	1		
B1	$i$	$ih_a^2$	1		$i$	$\frac{1}{2}ih_a^2$	1	
B2	$ih_a^2$	$i$	$h_b^2 + i^2h_a^2$	1	$\frac{1}{2}ih_a^2$	$i$	$\frac{1}{2}h_b^2 + \frac{1}{2}i^2h_a^2$	1

In the case where  $A \rightarrow B$  (A is AE and B is AE) the expected MZ and DZ correlation matrices would be:

	MZ twins				DZ twins			
	A1	A2	B1	B2	A1	A2	B1	B2
A1	1				1			
A2	$h_a^2 + i^2h_b^2$	1			$\frac{1}{2}h_a^2 + \frac{1}{2}i^2h_b^2$	1		
B1	$i$	$ih_b^2$	1		$i$	$\frac{1}{2}ih_b^2$	1	
B2	$ih_b^2$	$i$	$h_b^2$	1	$\frac{1}{2}ih_b^2$	$i$	$\frac{1}{2}h_b^2$	1.

If  $h_a = h_b$ ,  
then  
causation  
cannot be  
resolved

or practically.

**KEY WORDS:** Twins; reciprocal causation; genetics.

### INTRODUCTION

It is widely acknowledged that the existence of a correlation between two variables, measured at a single point in time, has no necessary implications about causation (Fisher, 1958). There are many ex-

the early environment has a direct causal influence on risk of psychopathology ( $B \rightarrow A$ ) or because current psychopathology is biasing recall of early experiences ( $A \rightarrow B$ ), or alternatively, both these processes may be operating simultaneously (recip-

Heath, Kessler, Neale, Hewitt, Eaves

proach doe  
ion and m  
d can be us  
d even wh  
for any me  
r time and  
informative

**Table IV.** Sample Sizes ( $N$  of Twin Pairs) Required for 80% Power of Rejecting False Unidirectional Hypothesis, for Selected Sets of Parameters Values of True Models<sup>a</sup>

	Trait A			Trait B			True model																
	$e^2$	$h^2$	$d^2$	$c^2$	$e'^2$	$h'^2$	$c'^2$	$A \rightarrow B$							$B \rightarrow A$								
								df	$i'=0.6$	$i'=0.5$	$i'=0.4$	$i'=0.3$	$i'=0.25$	$i'=0.2$	$i'=0.15$	df	$i=0.6$	$i=0.5$	$i=0.4$	$i=0.3$	$i=0.25$	$i=0.2$	$i=0.15$
								(Reject $B \rightarrow A$ )							(Reject $A \rightarrow B$ )								
I. No measurement error																							
1	0.25	0.5	0.25	0	0.25	0	0.75	1	93	114	157	248	337	503	861	2	111	141	195	313	432	652	1,127
	0.5	0.33	0.17	0	0.5	0	0.5	1	457	594	834	1,329	1,829	2,751	4,746	2	410	514	702	1,109	1,521	2,283	3,930
2	0.25	0.75	0	0	0.25	0	0.75	1	106	131	182	296	411	622	1,079	2	126	162	228	372	518	787	1,367
	0.5	0.5	0	0	0.5	0	0.5	1	521	687	992	1,655	2,321	3,547	6,200	2	485	617	862	1,308	1,828	2,822	5,080

What would be expect the Var-Cov  
Matrix to look like if  $V_1 \rightarrow V_2$

$$(I - B)^{-1} (A + C + E) (I - B)^{-1T}$$

$$\left( \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 0 & 0 & 0 & 0 \\ \beta & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \beta & 0 \end{bmatrix} \right)^{-1} \left( \begin{bmatrix} A_1 & 0 & 0 & 0 \\ 0 & A_2 & 0 & 0 \\ A_1 & 0 & A_1 & 0 \\ 0 & A_2 & 0 & A_2 \end{bmatrix} + \begin{bmatrix} C_1 & 0 & 0 & 0 \\ 0 & C_2 & 0 & 0 \\ C_1 & 0 & C_1 & 0 \\ 0 & C_2 & 0 & C_2 \end{bmatrix} + \begin{bmatrix} E_1 & 0 & 0 & 0 \\ 0 & E_2 & 0 & 0 \\ 0 & 0 & E_1 & 0 \\ 0 & 0 & 0 & E_2 \end{bmatrix} \right)$$

$$(I - B)^{-1} (A + C + E) (I - B)^{-1T}$$

Let's Assume:

$$V_1 \rightarrow V_2$$

$$\beta = .3$$

$V_1$  is ACE

$V_2$  is AE

$$\left( \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 0 & 0 & 0 & 0 \\ .3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & .3 & 0 \end{bmatrix} \right)^{-1}$$

$$\left( \begin{bmatrix} .33 & 0 & 0 & 0 \\ 0 & .50 & 0 & 0 \\ .33 & 0 & .33 & 0 \\ 0 & .50 & 0 & .50 \end{bmatrix} + \begin{bmatrix} .33 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ .33 & 0 & .33 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} .34 & 0 & 0 & 0 \\ 0 & .50 & 0 & 0 \\ 0 & 0 & .34 & 0 \\ 0 & 0 & 0 & .50 \end{bmatrix} \right)$$

## MZ Twins

1.00			
.30	1.09		
.66	.198	1.00	
.198	.559	.30	1.09

## DZ Twins

1.00			
.30	1.09		
.495	.149	1.00	
.149	.295	.30	1.09

$$V_1: .66/.495 = 1.333$$

$$V_1 | V_2: .198/.149 = 1.329$$

$$V_2: .559/.295 = 1.89$$

$$V_2 \rightarrow V_1$$

MZ Twins

DZ Twins

1.09

1.09

.30 1.00

.30 1.00

.705 .15 1.09

.518 .075 1.09

.15 .50 .30 1.00

.075 .250 .30 1.00

$$V_1: .705/.518 = 1.36$$

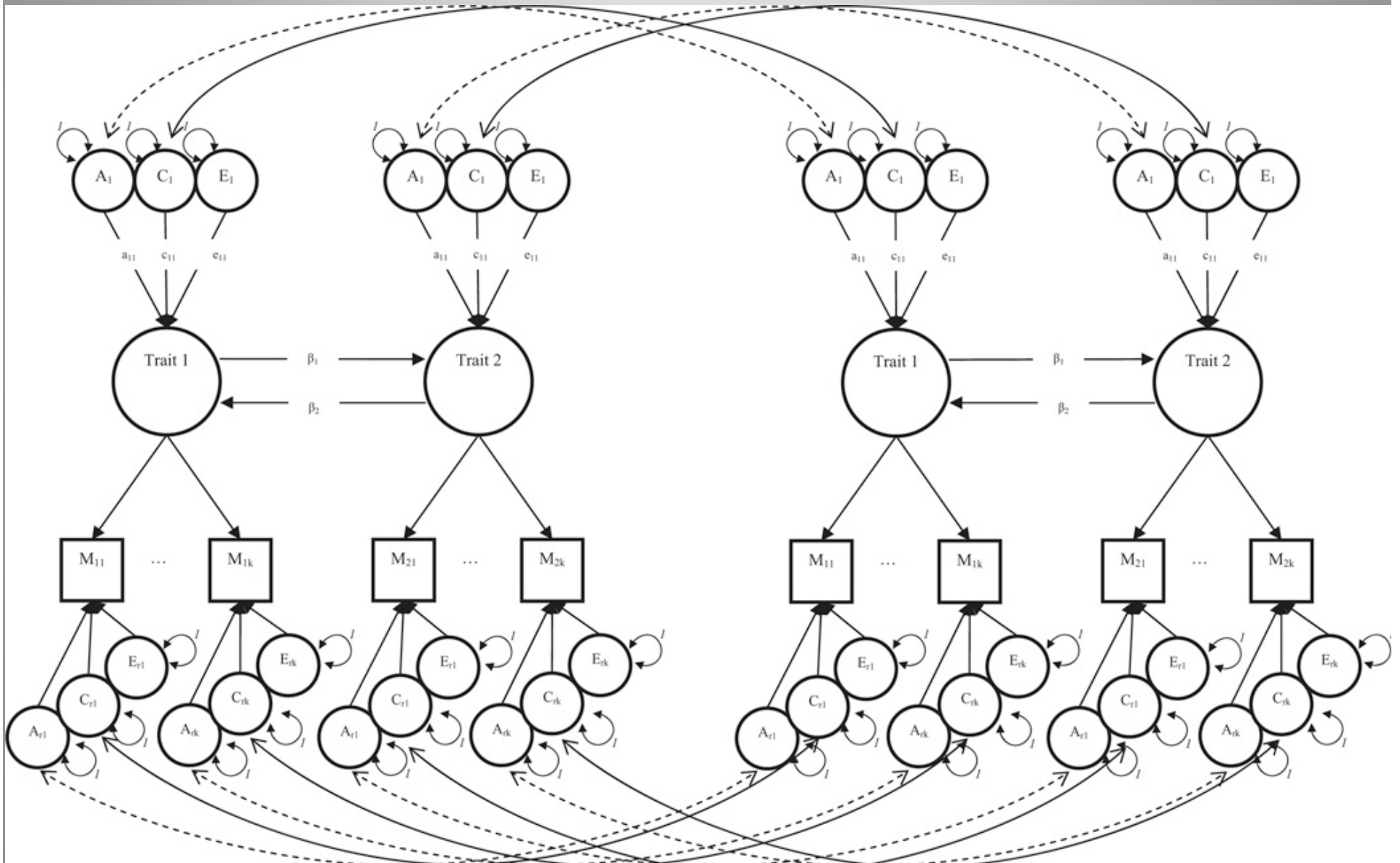
$$V_1 | V_2: .15/.075 = 2.00$$

$$V_2: .50/.25 = 2.00$$

# When is the Direction of Causation Model Appropriate

1. Strong hypothesis about the causal direction
2. No confounding intermediate variables
3. When the modes of transmission are distinct

# Where do things go wrong? Differential measurement error



## Testing Causality in the Association Between Regular Exercise and Symptoms of Anxiety and Depression

Marleen H. M. De Moor, MSc; Dorret I. Boomsma, PhD; Janine H. Stubbe; Gonneke Willemsen, PhD; Eco J. C. de Geus, PhD

**Conclusion:** Regular exercise is associated with reduced anxious and depressive symptoms in the population at large, but the association is not because of causal effects of exercise.



### Regular exercise, subjective wellbeing, and internalizing problems in adolescence: causality or genetic pleiotropy?

Meike Bartels<sup>1,2\*</sup>, Marleen H. M. de Moor<sup>1,2</sup>, Niels van der Aa<sup>1,2</sup>, Dorret I. Boomsma<sup>1,2</sup> and Eco J. C. de Geus<sup>1,2</sup>

<sup>1</sup> Department of Biological Psychology, VU University Amsterdam, Amsterdam, Netherlands

<sup>2</sup> EMGO Institute for Health and Care, VU University Medical Centre, Amsterdam, Netherlands

sectionally and longitudinally. We conclude that exercise behavior is associated with fewer internalizing problems and higher levels of SWB. The association largely reflects the effects of common genetic factors on these traits.

## A Twin-Sibling Study on the Relationship Between Exercise Attitudes and Exercise Behavior

Charlotte Huppertz · Meike Bartels · Iris E. Jansen · Dorret I. Boomsma · Gonneke Willemsen · Marleen H. M. de Moor · Eco J. C. de Geus

present as well. Furthermore, after taking genetic pleiotropy into account, our data were compatible with a causal association between exercise attitudes and exercise behavior. Replication in longitudinal studies is now needed to more firmly establish this causality and its direction.