A Woefully Incomplete History of Empirical Applications of Algorithmic Causal Discovery (ACD)

UAI 2023

Workshop:

The History and Development of Search Methods for Causal Structure

Richard Scheines

Carnegie Mellon University

Literature searches

Thanks to:

Yujia Zheng

Ph.D. Student

- 2013 Workshop at CMU on empirical applications of ACD:
  https://www.cmu.edu/dietrich/philosophy/events/workshops-conferences/causal-discovery/index.html

- Work since by authors from Workshop at CMU:
  https://docs.google.com/document/d/1nv84oQYy9YAXvn5RxNeGcaGwsmv_RlLKE50hHNWWxcw/edit

- Other Published Applications of “Causal Discovery” Methods 2003 – 2023:

https://docs.google.com/document/d/1mcjLIFk8JRUVN99rfv3BGAPUSoCGtC13ZiY1pY4wUMk/edit?usp=sharing

The History of Applying ACD to Empirical Data

- ACD tools ~ 40 years : “Structure Learning”, “Causal Discovery”, etc.
- Now: dozens of algorithms, toolkits, full suites available, in Java, Python, R, etc.
- Massive # of algorithm + simulation studies / small # empirical applications – why?
- ACD requires assumptions, e.g., about:
  - Time
  - Feedback
  - Distributions
  - Parametric families
  - Sampling
  - Latent confounders
  - Measurement / measurement error

Still – there are 100s of applications –

How to categorize them for this presentation?

- Algorithm
- Assumptions
- Discipline
- Experimental follow-up
- Scientific purpose

Common Scientific Purposes of Using of ACD on Empirical Data

1. Finding alternatives …. to:
    1. Theoretically specified causal hypotheses
    2. Regression based causal inferences / standard factor analysis latent variable models
2. Generating candidate causes
3. Mechanism Discovery

Presentation goal: see the rough outlines of the landscape,

So: simple and shallow and fast

Finding alternatives …. to:

1. “Theoretically” specified causal hypotheses

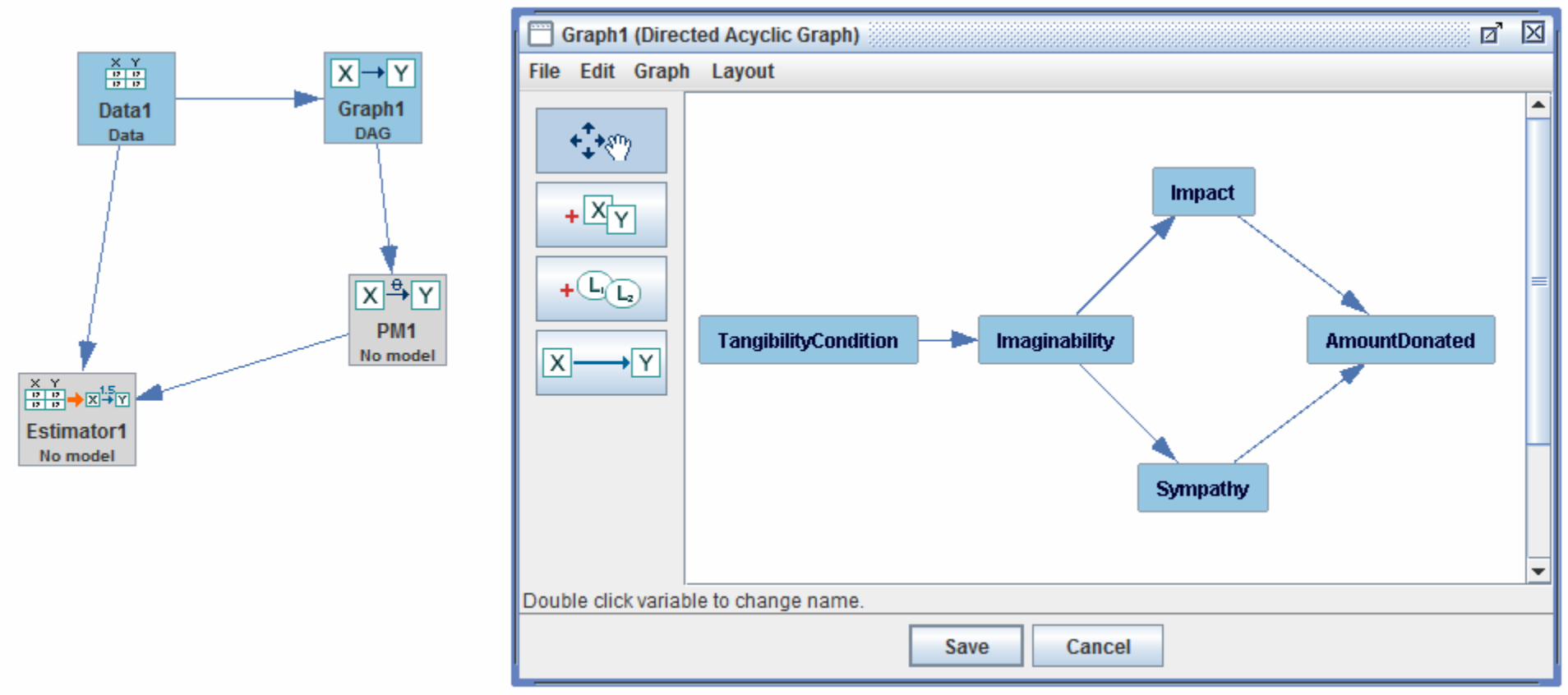
Charitable Giving

What influences giving? Specificity? Sympathy? Impact?

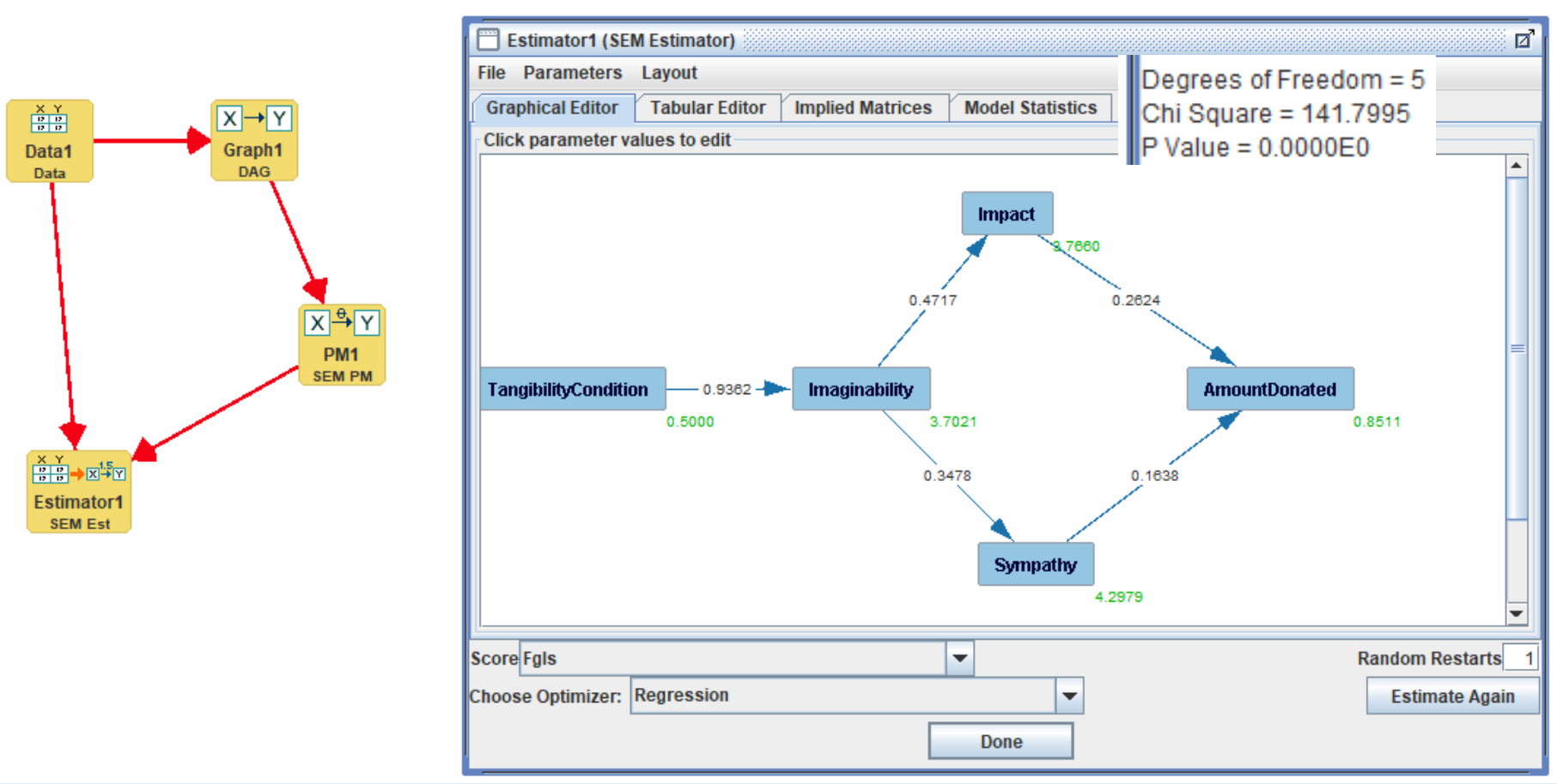
"The Donor is in the Details", (2013) Organizational Behavior and Human Decision Processes, Issue 1, 15-23, C. Cryder, with G. Loewenstein, R. Scheines.

		N = 94
TangibilityCondition	[1,0]	Randomly assigned experimental condition
Imaginability	[1..7]	How concrete is the scenario
Sympathy	[1..7]	How much sympathy for target
Impact	[1..7]	How much impact will my donation have
AmountDonated	[0..5]	How much actually donated

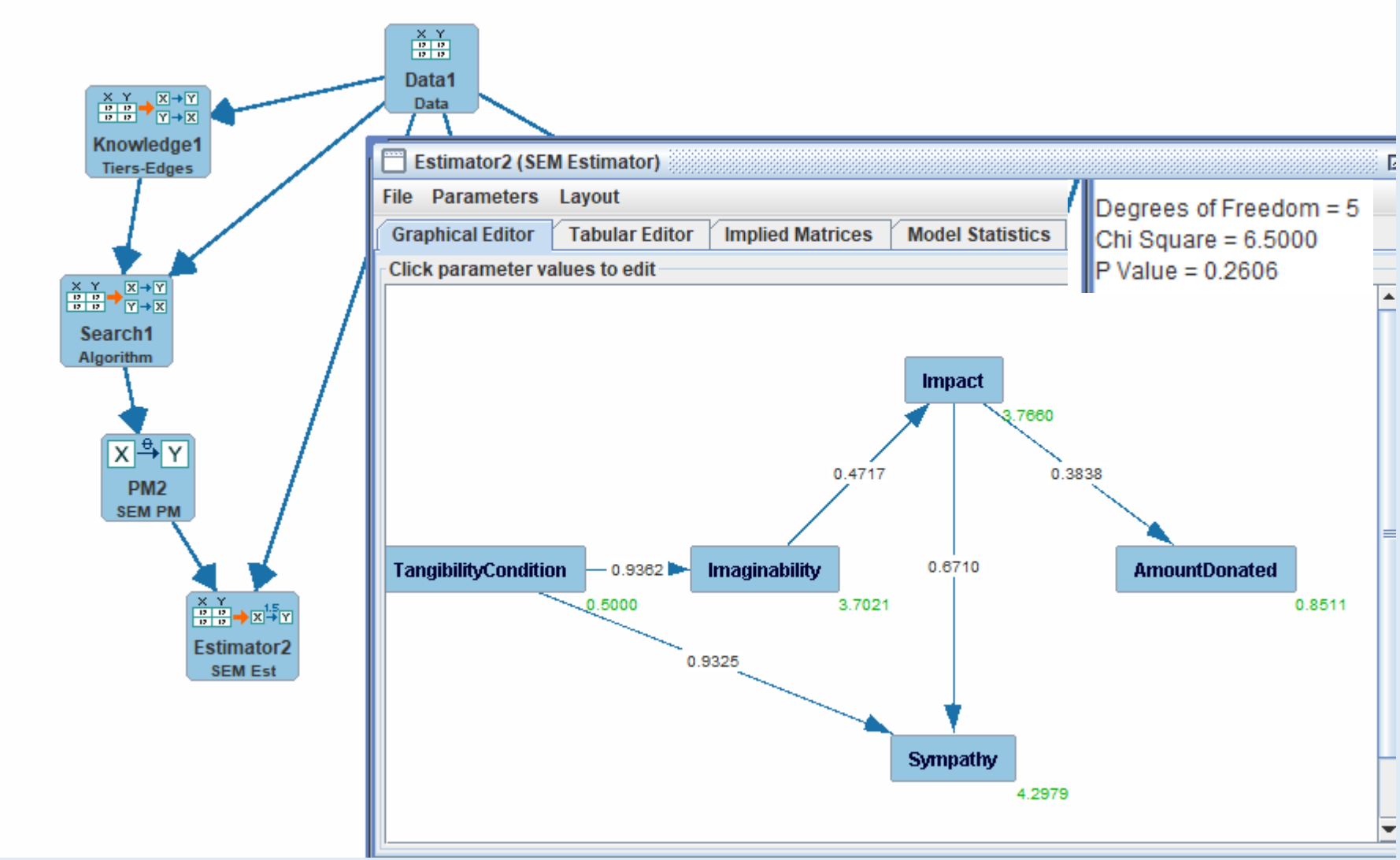
Theoretical Hypothesis



Estimated Model



Search Alternative



Common Scientific Purposes of Using of ACD on Empirical Data

1. Finding alternatives …. to:
    1. Theoretically specified causal hypotheses
    2. Regression based causal inferences / standard factor analysis latent variable models

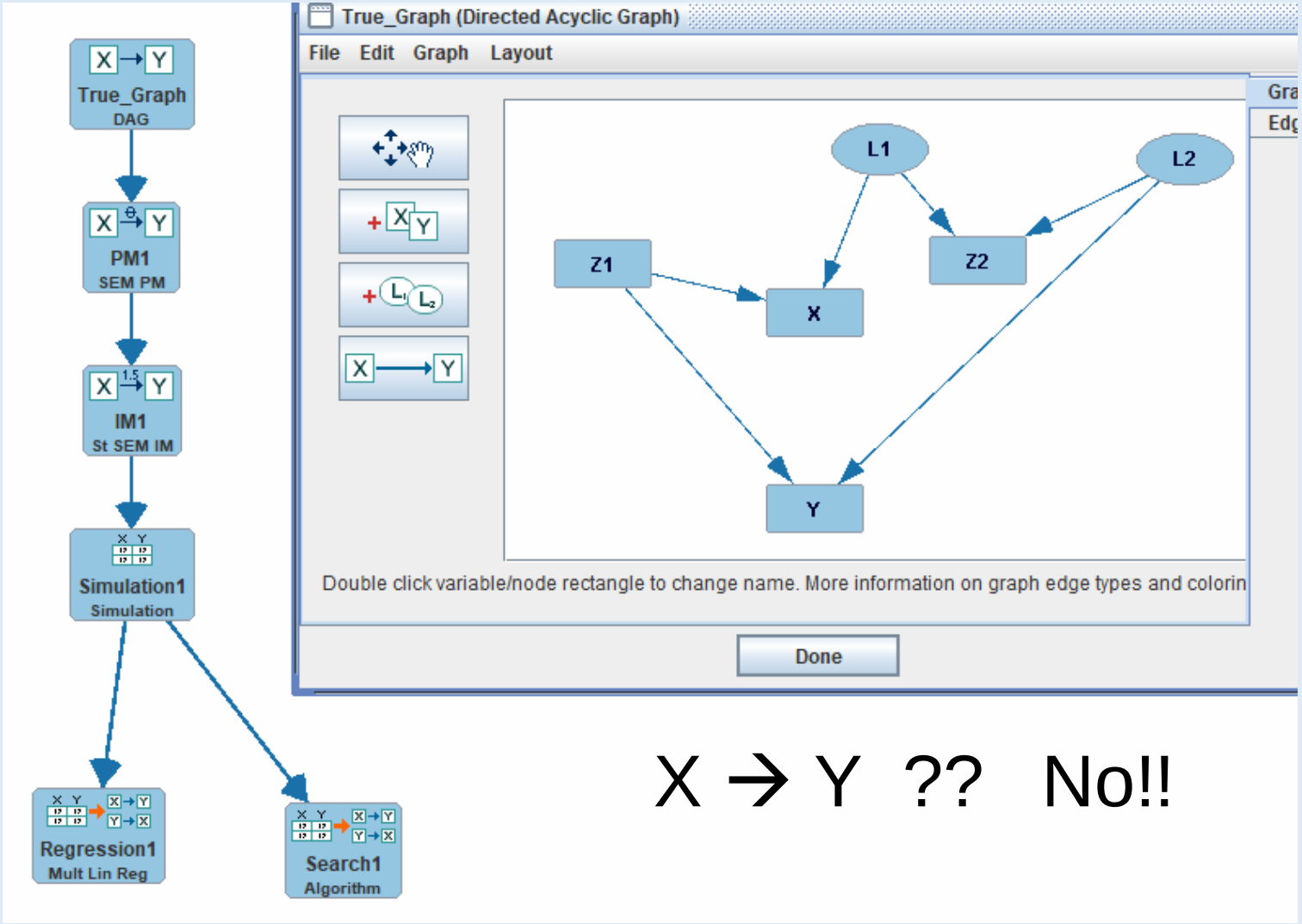
Regression for Causal Inference

- X a prima facie cause of Y? (i) plausible, and (ii) X and Y associated

- Regress Y on X and potential confounders Z_i:
  - Z_i associated with X, and
  - Z_i associated with Y, and
  - Z_i prior to X

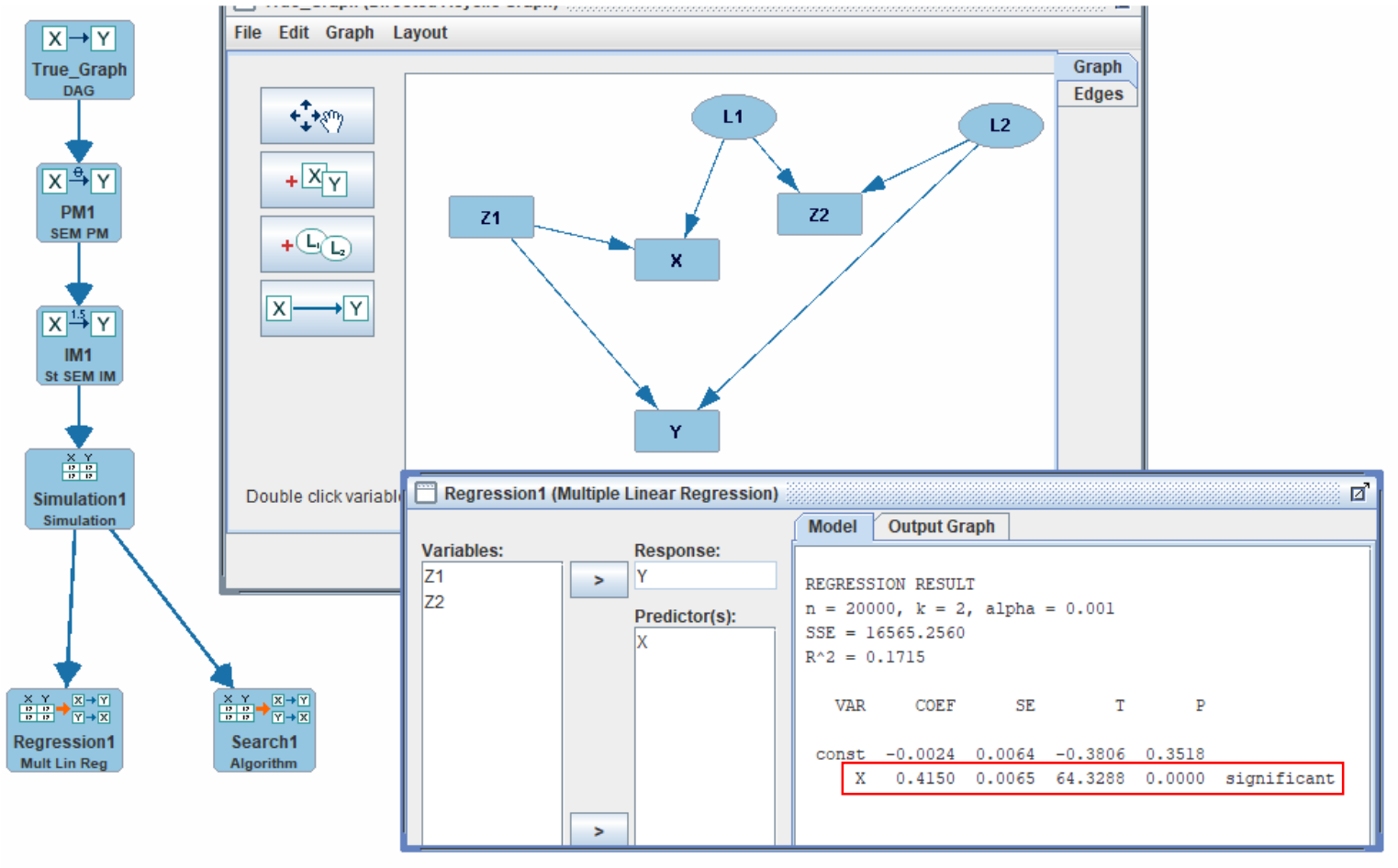
Regression is bad for causal inference – ACD is better

No practicing social scientist or statistician seems to know this – or take it seriously



Prima facie cause? Yes

N = 20,000

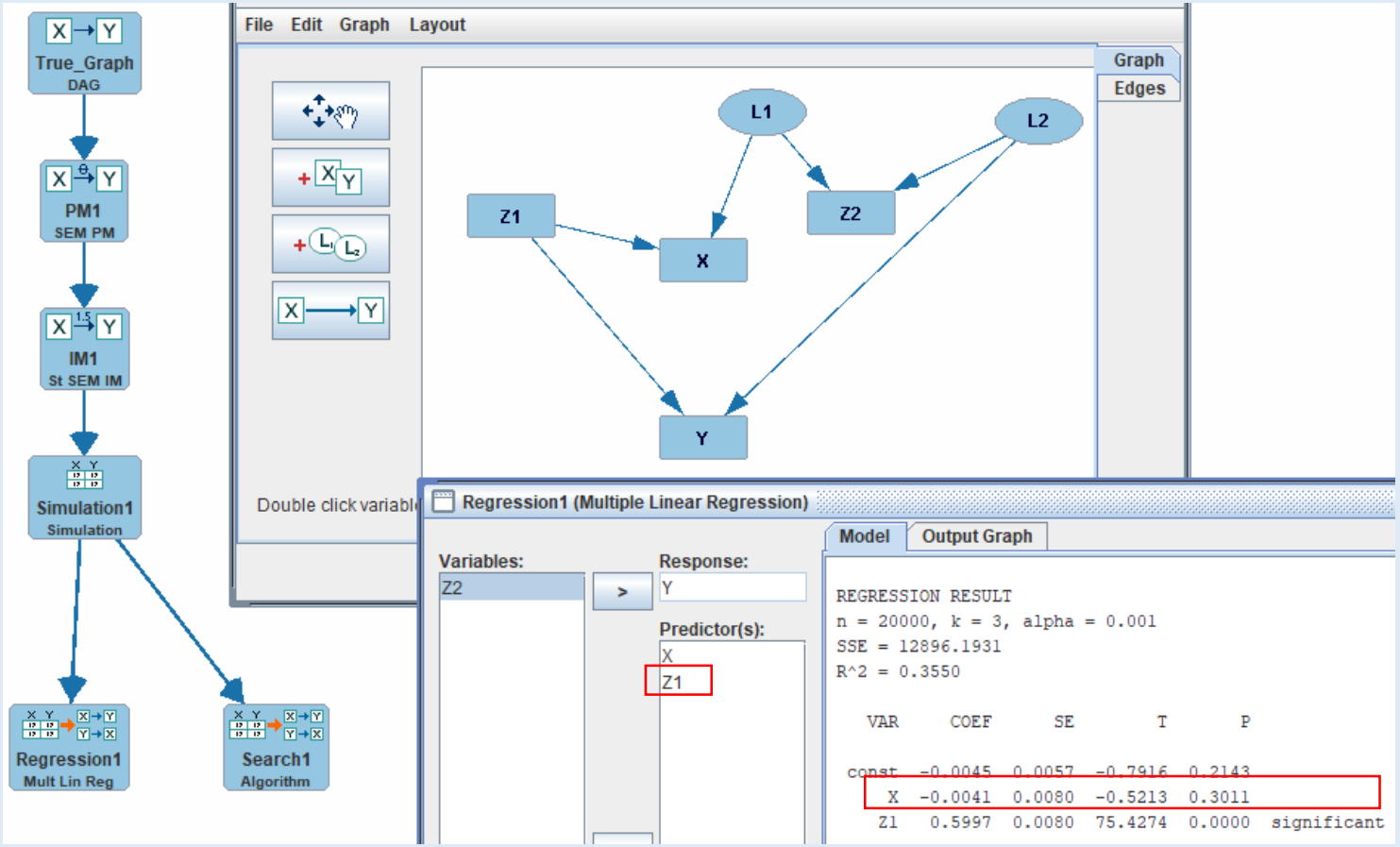


X → Y ??

Confounders to control for in a regression?

Z1 and Z2

N = 20,000

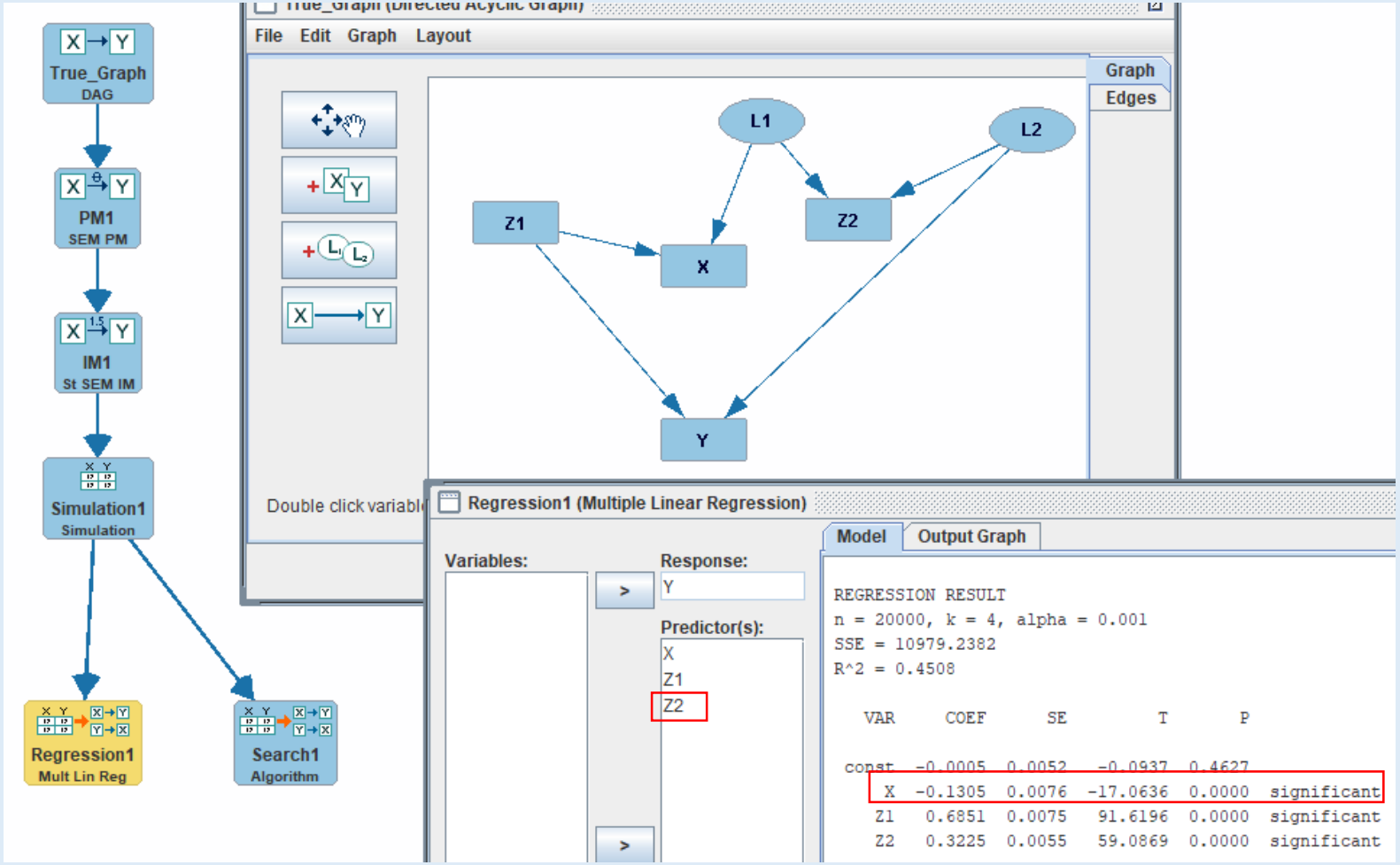


X → Y ??

X → Y ??

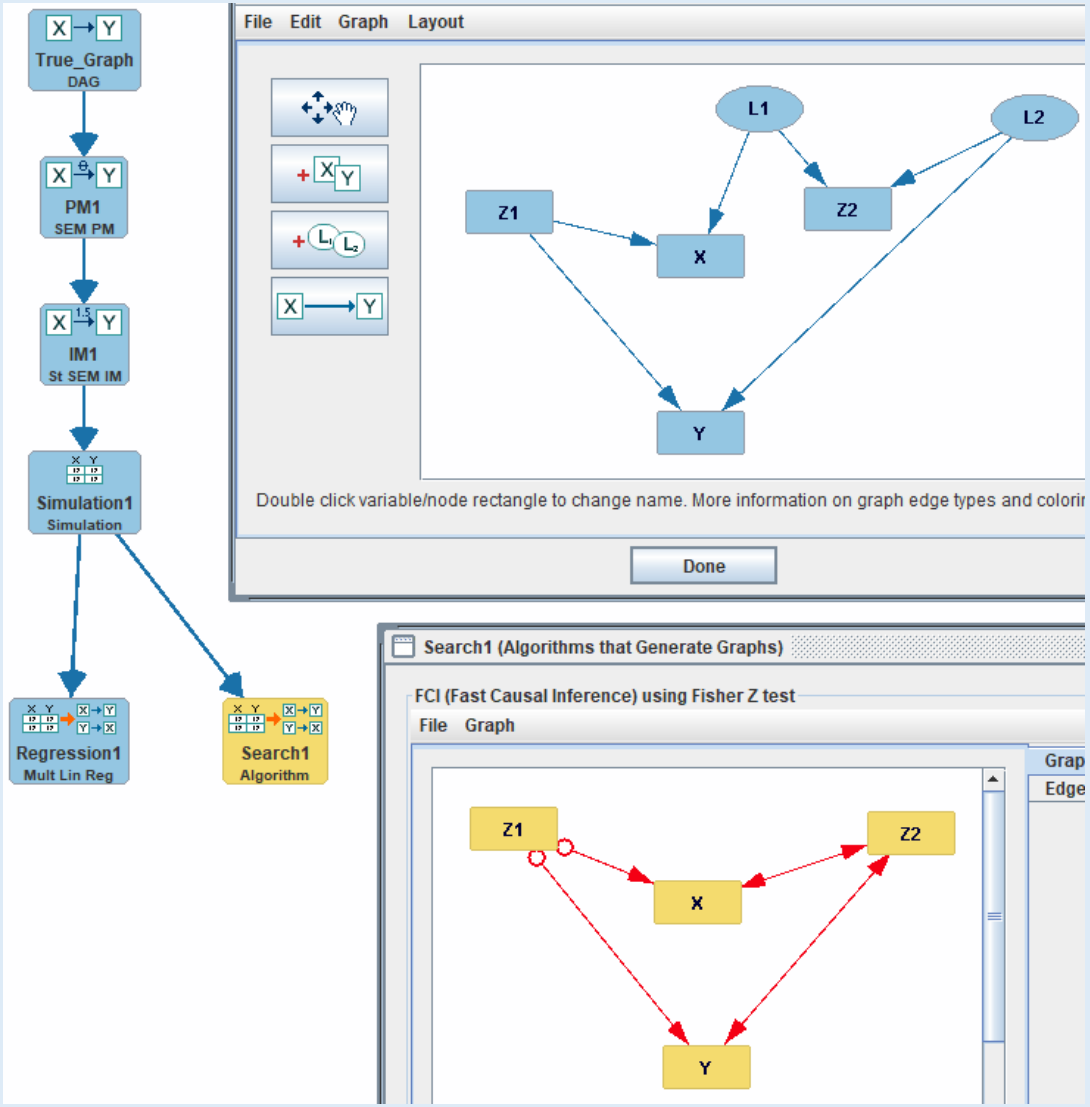
Regression: Yes - *incorrect*!!

N = 20,000



X → Y ?? ACD: No - correct!!

N = 20,000



X-Z2 confounded - correct!!

Z2-Y confounded - correct!!

Z2 a collider, must *not* be controlled for? correct!!

Z1 a non-collider, must be controlled for? correct!!

Foreign Investment

Does Foreign Investment in 3rd World Countries cause Repression?

Timberlake, M. and Williams, K. (1984). Dependence, political exclusion, and government repression: Some cross-national evidence. *American Sociological Review* 49, 141-146.

N = 72

po	degree of political exclusivity
cv	lack of civil liberties
en	energy consumption per capita (economic development)
fi	level of foreign investment

Foreign Investment

Correlations

	po	fi	en
fi	-.175		
en	-.480	0.330	
cv	0.868	-.391	-.430

Prima Facie:

Foreign Investment associated with LESS Political exclusion
(more Political repression)

Foreign Investment

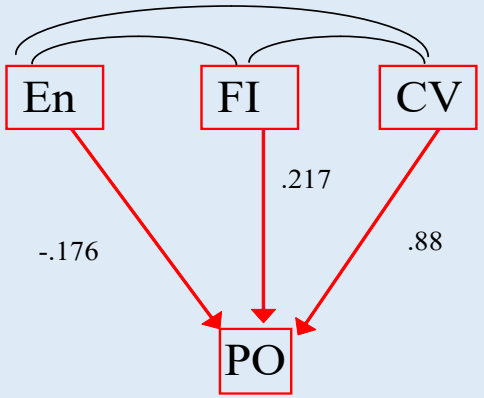
Regression Results

po =	.227*fi	- .176*en +	.880*cv	
SE	(.058)		(.059)	(.060)
t	3.941		-2.99	14.6

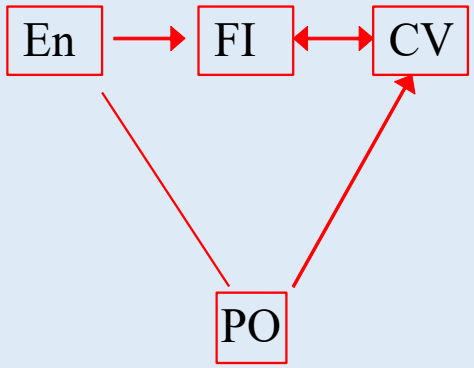
Controlling for Economic Development (en) and lack of civil liberties (cv) flips the sign:

Foreign Investment causes More Political exclusion
(LESS Political repression)

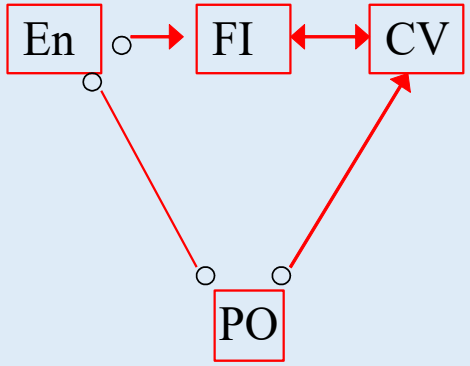
Case Study: Foreign Investment *Alternative Models*



Regression

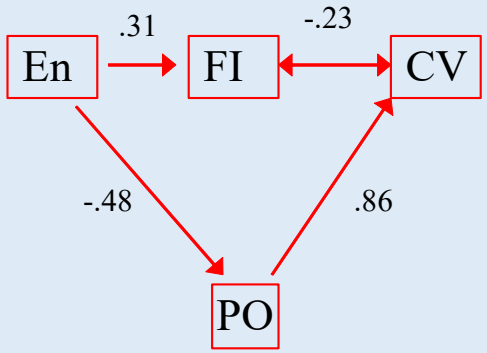


Tetrad - PC



Tetrad - FCI

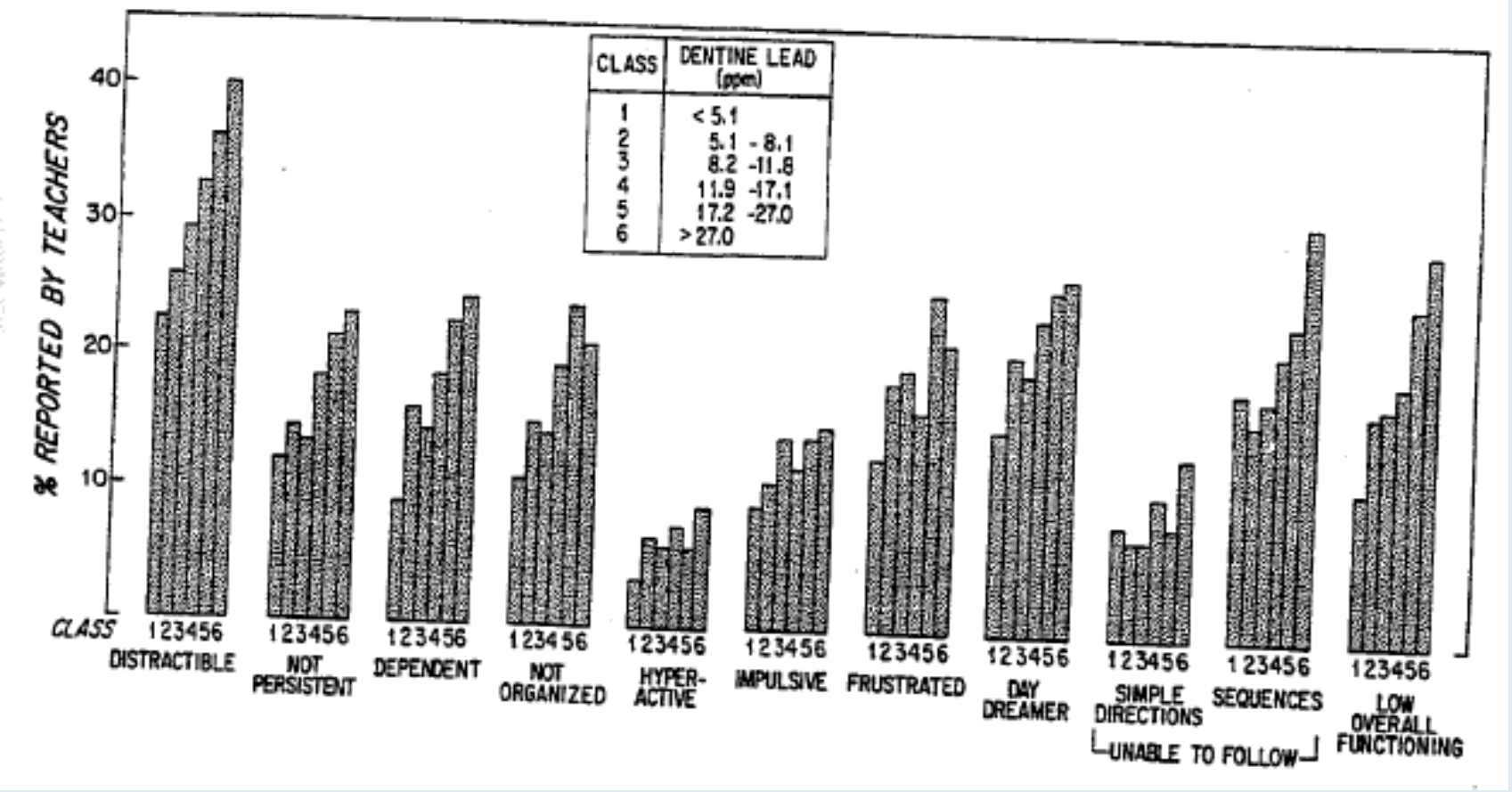
There is no linear model with testable constraints (df > 0) in which FI has a positive effect on PO, that is not rejected by the data.



Fit: df=2, $\chi 2$=0.12,
p-value = .94

Lead → IQ 1979 NEJM Study

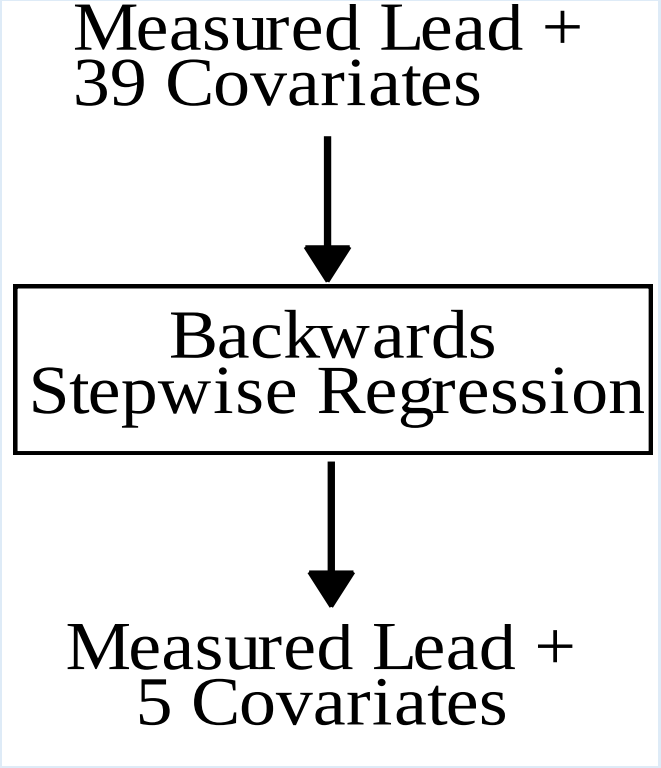
- 2500 children's teeth collected and measured for lead exposure
- All children rated behaviorally by teacher



1979 NEJM Study

- 39 Potential confounding variables measured:
  - Socioeconomic status
  - Parental IQ
  - Mother’s age at birth
  - Father’s age at birth
  - Mother’s level of education
  - Number of live birth’s before sampled child
  - Etc.

Lead and IQ: Variable Selection



Final Variables (Needleman)

-lead baby teeth

-fab father's age

-mab mother's age

-nlb number of live births

-med mother's education

-piq parent's IQ

-ciq child's IQ

Needleman Regression

- standardized coefficient
- (t-ratios in parentheses)
- p-value for significance

ciq = - .143 lead - .204 fab - .159 nlb + .219 med + .237 mab + .247 piq

(2.32) (1.79) (2.30) (3.08) (1.97) (3.87)

0.02 0.09 0.02 <0.01 0.05 <0.01

All variables significant at .1 R^2 = .271

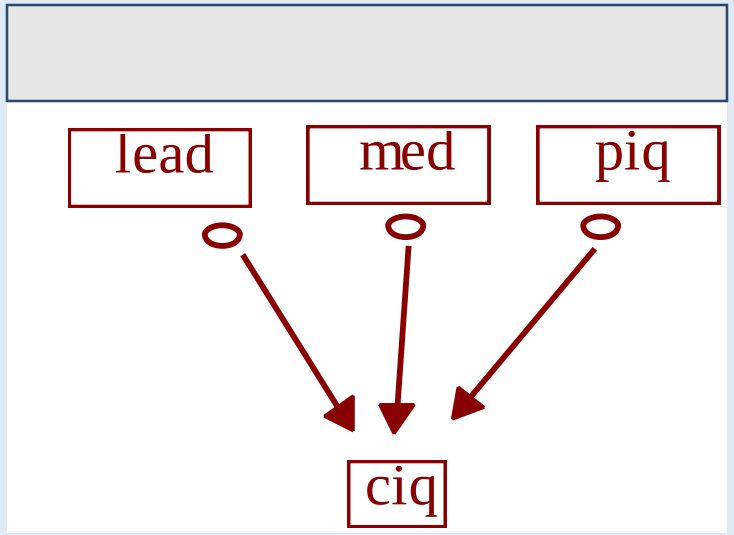
TETRAD Variable Selection

Tetrad

mab _||_ ciq

fab _||_ ciq

nlb _||_ ciq | med



Regressions

- standardized coefficient
- (t-ratios in parentheses)
- p-value for significance

Needleman (R^2 = .271)

ciq =	- .143 lead	- .204 fab	- .159 nlb	+ .219 med	+ .237 mab	+ .247 piq
	(2.32)	(1.79)	(2.30)	(3.08)	(1.97)	(3.87)
	0.02	0.09	0.02	<0.01	0.05	<0.01

TETRAD (R^2 = .243)

ciq =	- .177 lead	+ .251 med +	.253 piq
	(2.89)	(3.50)	(3.59)
	<0.01	<0.01	<0.01

Common Scientific Purposes of Using of ACD on Empirical Data

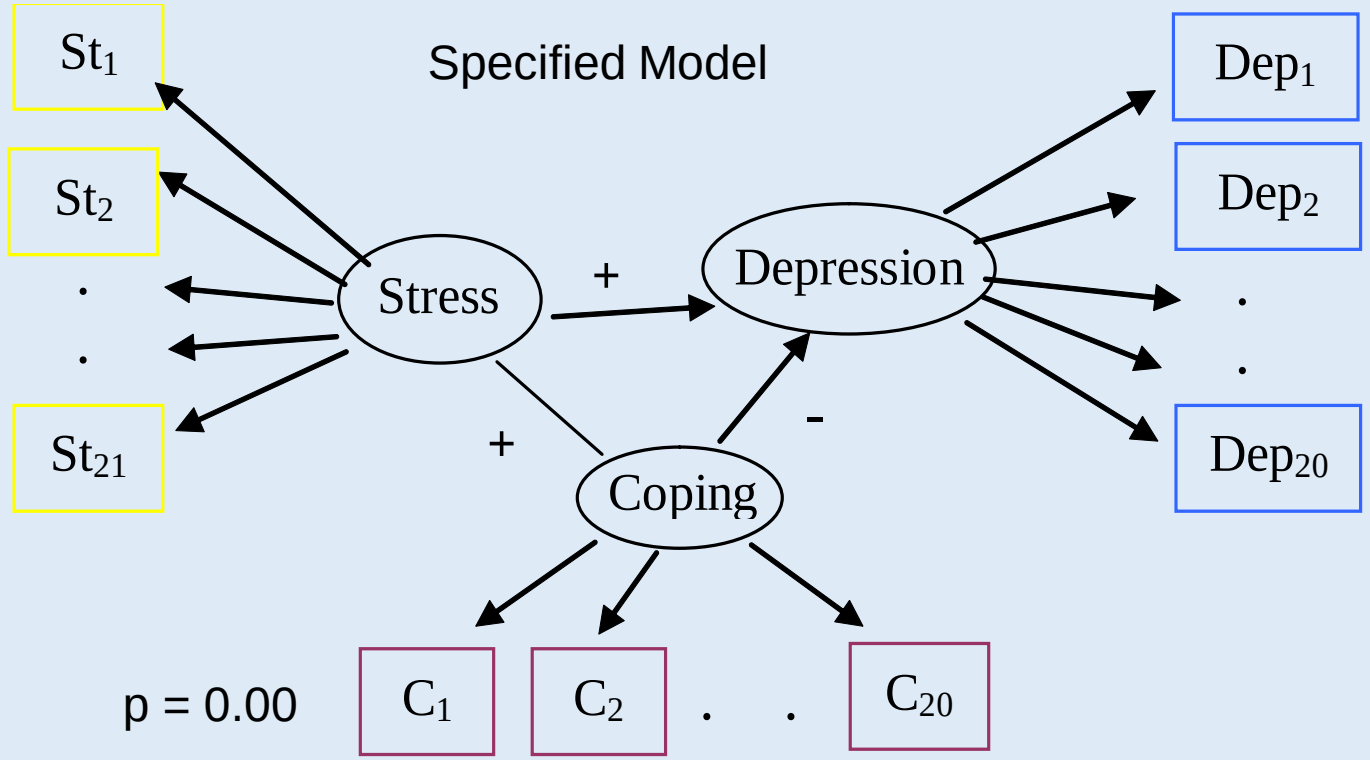
1. Finding alternatives …. to:
    1. Theoretically specified causal hypotheses
    2. Regression based causal inferences /

    standard factor analysis latent variable models

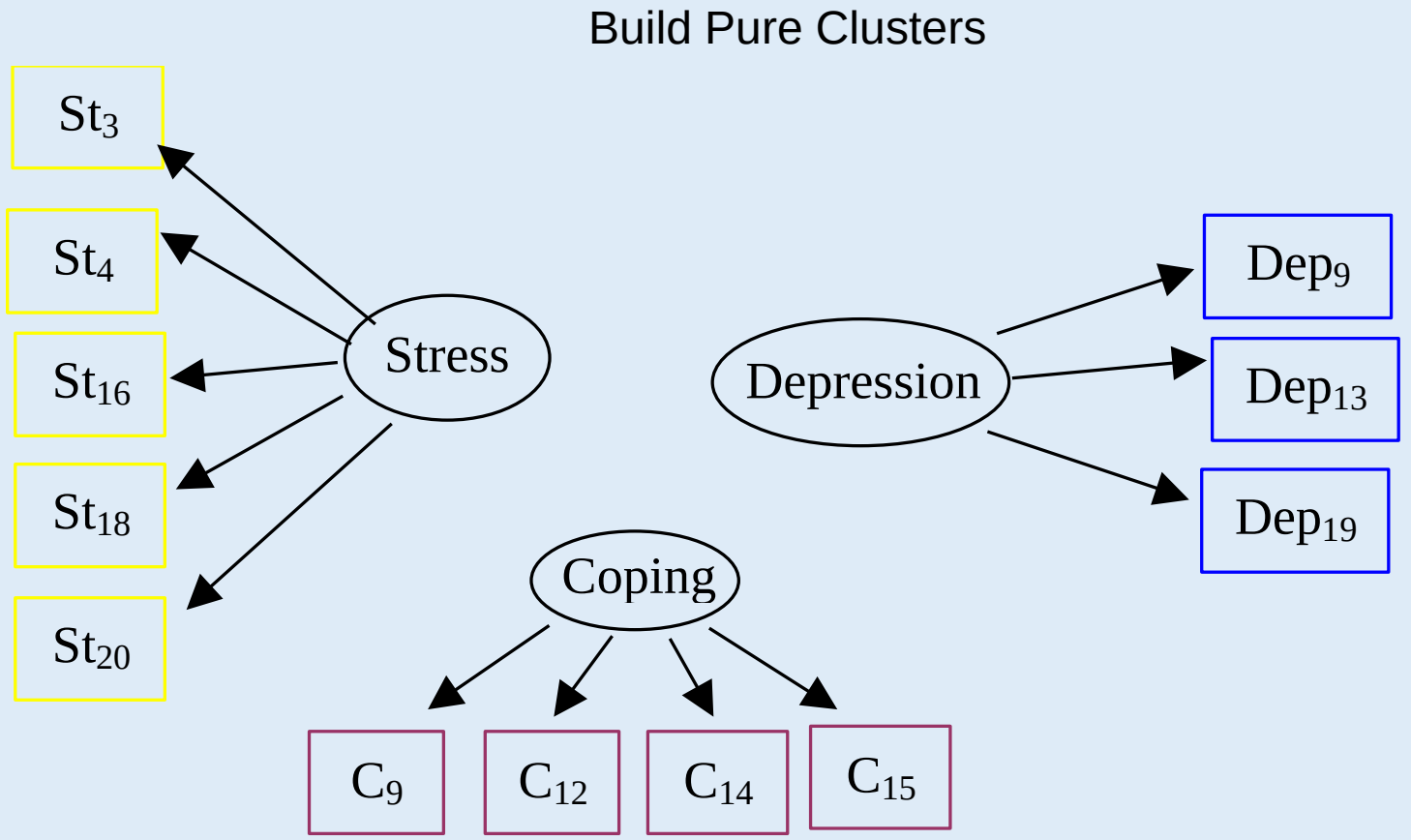
Case Study: Stress, Depression, and Religion

MSW Students (N = 127) 61 - item survey (Likert Scale)

- Stress: St_1 - St_{21}
- Depression: D_1 - D_{20}
- Religious Coping: C_1 - C_{20}

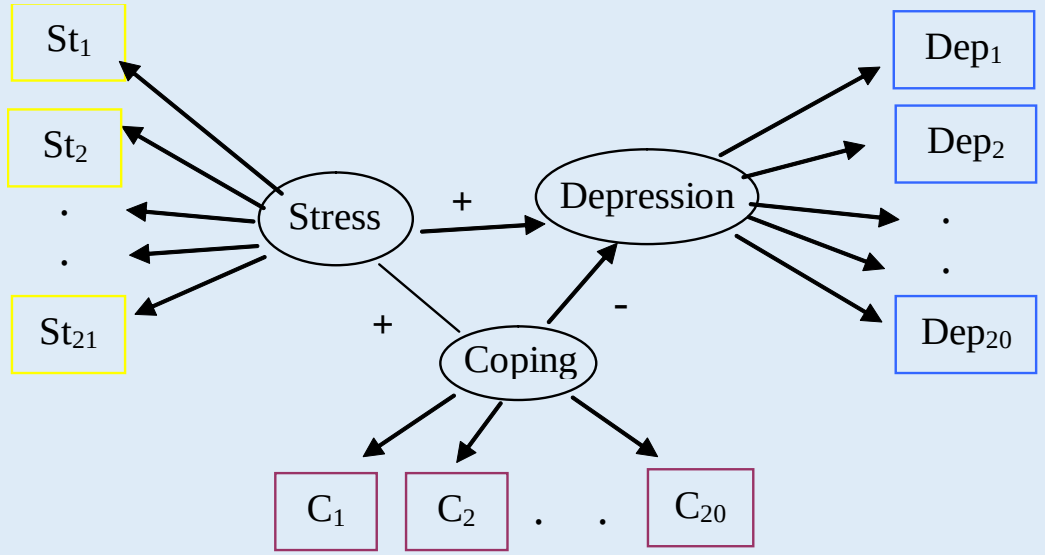


Case Study: Stress, Depression, and Religion



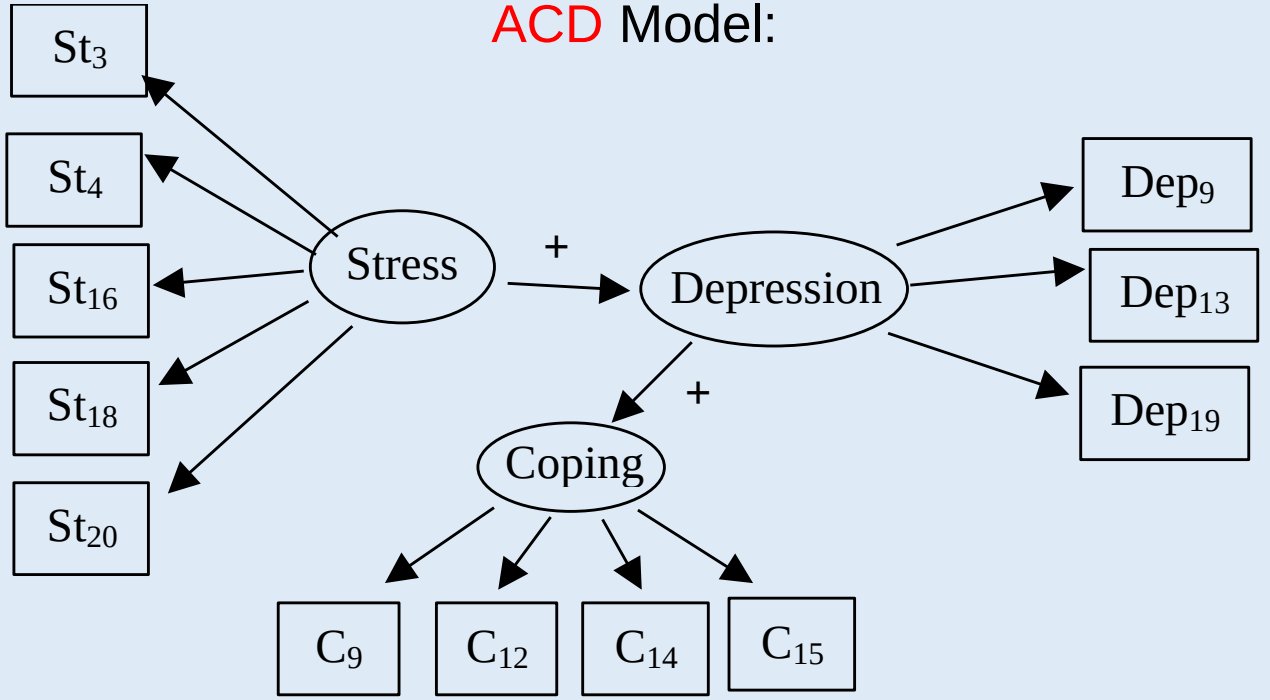
Case Study: Stress, Depression, and Religion

Specified Model



p = 0.00

ACD Model:



p = 0.28

Common Scientific Purposes of Using of ACD on Empirical Data

1. Finding alternatives …. to:
    1. Theoretically specified causal hypotheses
    2. Regression based causal inferences / standard factor analysis latent variable models

2. Generating candidate causes

3. Mechanism Discovery

Candidate Gene Regulators

Which genes regulate flowering time in *Arabidopsis thaliana*?

* Stekhoven DJ, et al. Causal stability ranking. *Bioinformatics* 28 (2012) 2819-2823.

Observational Data

- *n* = 47 *Arabidopsis thaliana* gene expression profiles of 4-day old seedlings for which subsequent flowering time was also measured

- Affymetrix ATH1 arrays with expression measurements on 21,440 *A. thaliana* genes

Candidate Gene Selection

Observational data:
Microarray measurements (**X**)
and flowering time (Y)

Causal Analysis

Causal networks on the variables in X and Y

Select those genes that reliably cause the most substantial changes in flowering time (Y) according to the causal networks

Candidate gene regulators of flowering time

Candidate Regulators of Flowering Time

- Considered the 25 genes that were ranked most likely to be causes of flowering time,
  according to the causal network analysis

- 5 of those 25 genes were known regulators of flowering

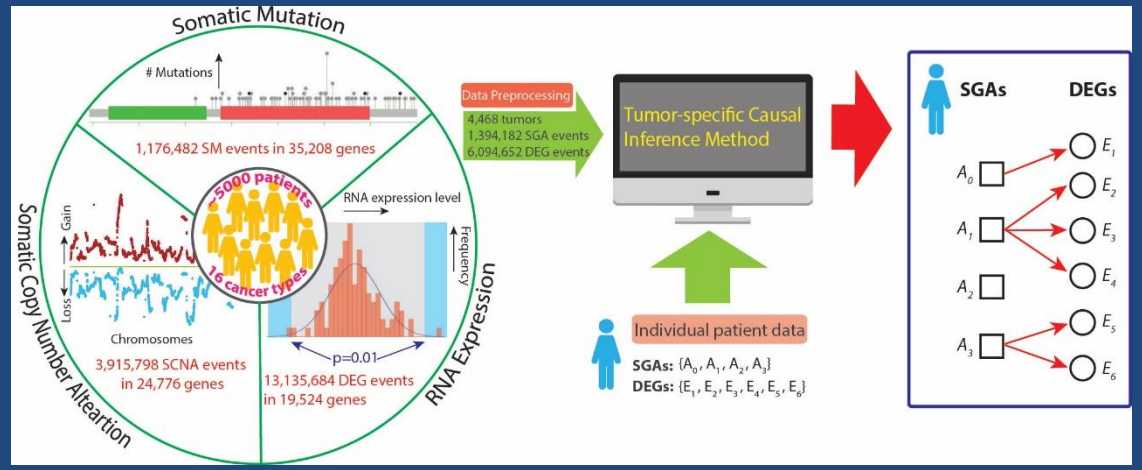
- 13 of those 25 genes were not known regulators and mutant seeds for each of them were available

Experimental Details

Greenhouse experiments on flowering time

- Seeds were planted, and flowering time was measured in days to bolting
- Seeds types yielding 4 or more plants were considered viable for analysis (9)
- 4 of 9 had a statistically significantly shorter mean flowering time ($p < 0.05$) than the control, wild-type plants
- Correlational and other analyses to identify candidate genes were successful roughly at chance rates, i.e., they were worthless

The Tumor-specific Driver Identification (TDI) Algorithm



PLoS Comput Biol. 2019 Jul 5;15(7):e1007088. doi: 10.1371/journal.pcbi.1007088. eCollection 2019 Jul.

Systematic discovery of the functional impact of somatic genome alterations in individual tumors through tumor-specific causal inference

Chunhui Cai [1 2], Gregory F Cooper [1 2], Kevin N Lu [1 2], Xiaojun Ma [1], Shuping Xu [3], Zhenlong Zhao [3], Xueer Chen [1 2], Yifan Xue [1 2], Adrian V Lee [2 3 4 5], Nathan Clark [2 6], Vicky Chen [1 2], Songjian Lu [1 2], Lujia Chen [1 2], Liyue Yu [1 2], Harry S Hochheiser [1 2], Xia Jiang [1 2], Q Jane Wang [3], Xinghua Lu [1 2 5]

Affiliations + expand

PMID: 31276486 PMCID: PMC6650088 DOI: 10.1371/journal.pcbi.1007088

Free PMC article

Abstract

Cancer is mainly caused by somatic genome alterations (SGAs). Precision oncology involves identifying and targeting tumor-specific aberrations resulting from causative SGAs. We developed a novel tumor-specific computational framework that finds the likely causative SGAs in an individual tumor and estimates their impact on oncogenic processes, which suggests the disease mechanisms that are acting in that tumor. This information can be used to guide precision oncology. We report a tumor-specific causal inference (TCI) framework, which estimates causative SGAs by modeling causal relationships between SGAs and molecular phenotypes (e.g., transcriptomic, proteomic, or metabolomic changes) within an individual tumor. We applied the TCI algorithm to tumors from The Cancer Genome Atlas (TCGA) and estimated for each tumor the SGAs that causally regulate the differentially expressed genes (DEGs) in that tumor. Overall, TCI identified 634 SGAs that are predicted to cause cancer-related DEGs in a significant number of tumors, including most of the previously known drivers and many novel candidate cancer drivers. The inferred causal relationships are statistically robust and biologically sensible, and multiple lines of experimental evidence support the predicted functional impact of both the well-known and the novel candidate drivers that are predicted by TCI. TCI provides a unified framework that integrates multiple types of SGAs and molecular phenotypes to estimate which genome perturbations are causally influencing one or more molecular/cellular phenotypes in an individual tumor. By identifying major candidate drivers and revealing their functional impact in an individual tumor, TCI sheds light on the disease mechanisms of that tumor, which can serve to advance our basic knowledge of cancer biology and to support precision oncology that provides tailored treatment of individual tumors.

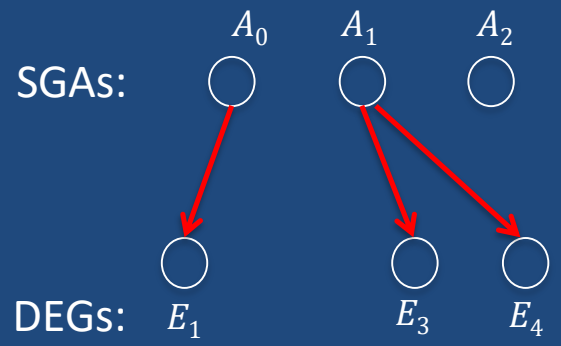
Identifying Tumor-specific Drivers

- Goal: identifying driver somatic genomic alterations (SGAs) of individual tumors

- A tumor usually hosts hundreds to thousands SGAs
  - Many passengers
  - Relatively few drivers

- Current knowledge of cancer driver genes is incomplete
  - 10 to 20% of tumors have no known drivers
  - 50% of tumors have ≤ 3 known drivers

Tumor-specific Driver Identification (TDI) Algorithm*



- Uses a bipartite graph representation
- Searches the graph for somatic genomic alterations (SGAs) that account for differentially expressed genes (DEGs) involved in an oncogenic process.
- Uses a Bayesian evaluation measure, which is tumor specific

Hypothesis: Those SGAs that predict DEGs well are good candidates as drivers of cancer

Results: Many known drivers recovered, several new candidate drivers discovered, many experimentally verified (6/6 perturbation confirmed).

Common Scientific Purposes of Using of ACD on Empirical Data

1. Finding alternatives …. to:
    1. Theoretically specified causal hypotheses
    2. Regression based causal inferences / standard factor analysis latent variable models

2. Generating candidate causes

3. Mechanism Discovery

Mechanism Discovery

1. Educational Research
2. Brain Research
3. Climate
4. Miscellany (lots!)
5. Economics (for Kevin Hoover)

Educational Research

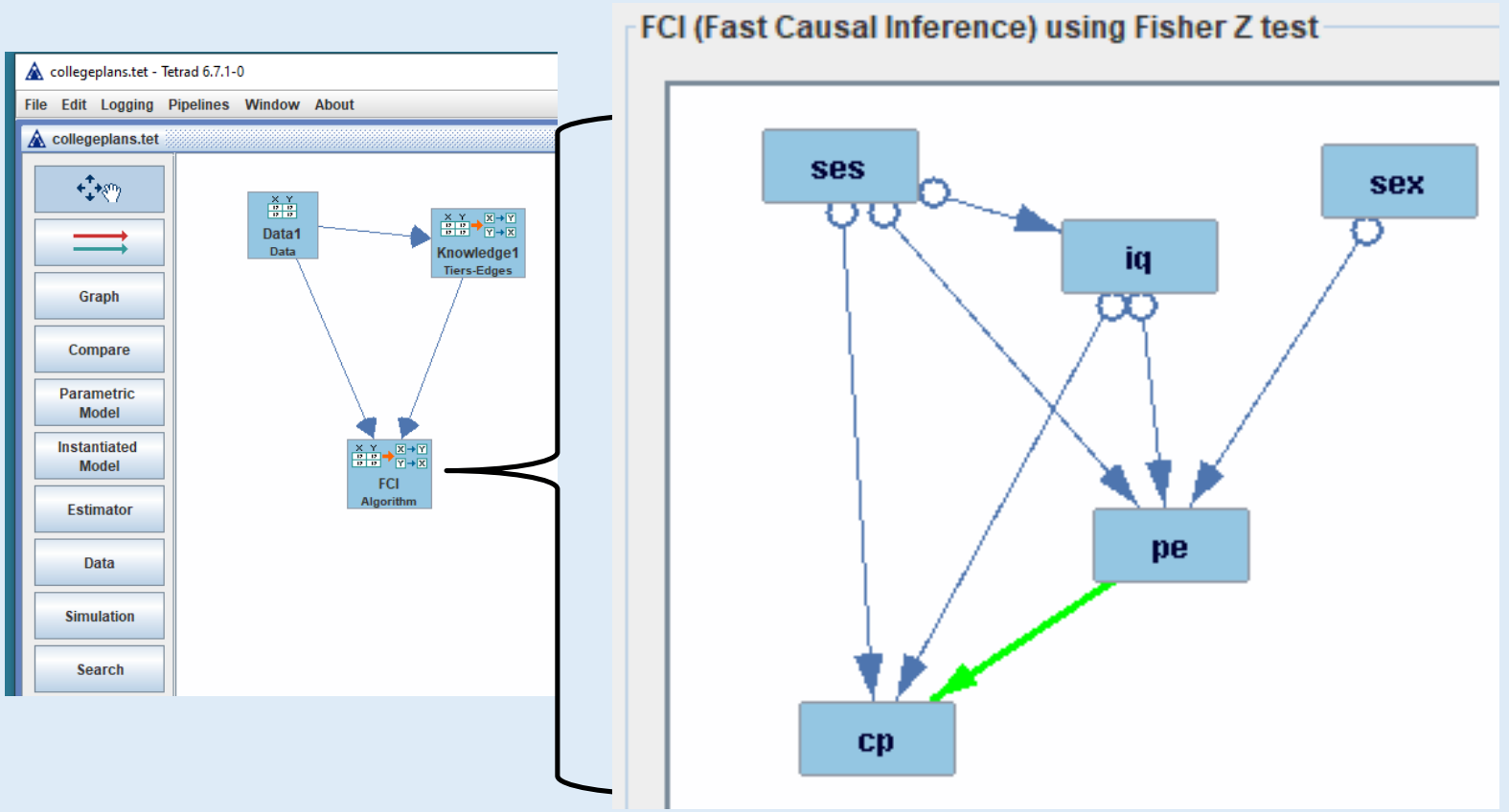
College Plans - Wisconsin HS seniors, 1979, (N > 10,000)

Questions:

1. Does parental encouragement depend on sex, even controlling for SES and IQ
2. Does parental encouragement matter?

Measures:

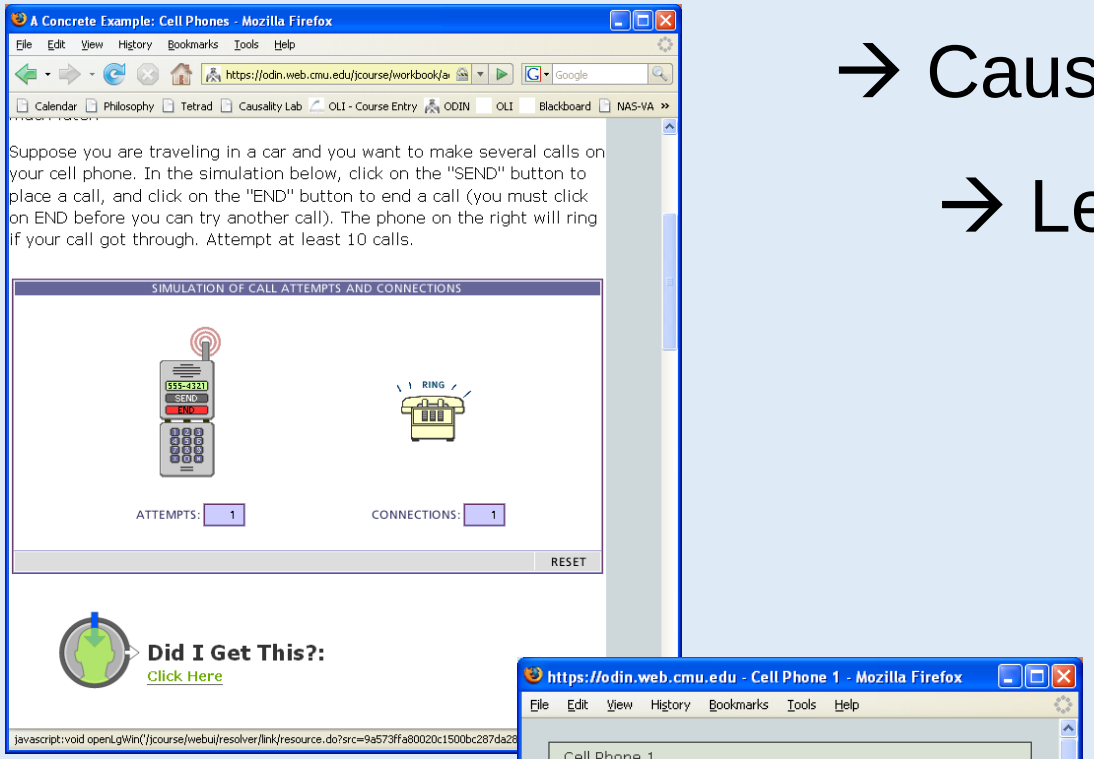
- SES
- IQ
- SEX
- PE (parental encouragement
- CP (College Plans)

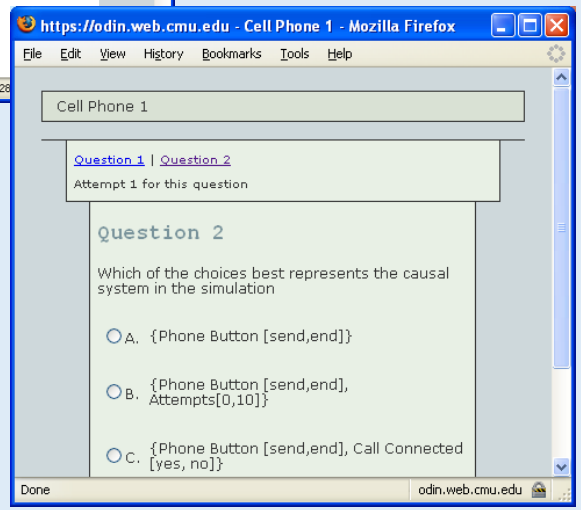


Online Course Data

→ Causal Discovery Methods

→ Learning Mechanism



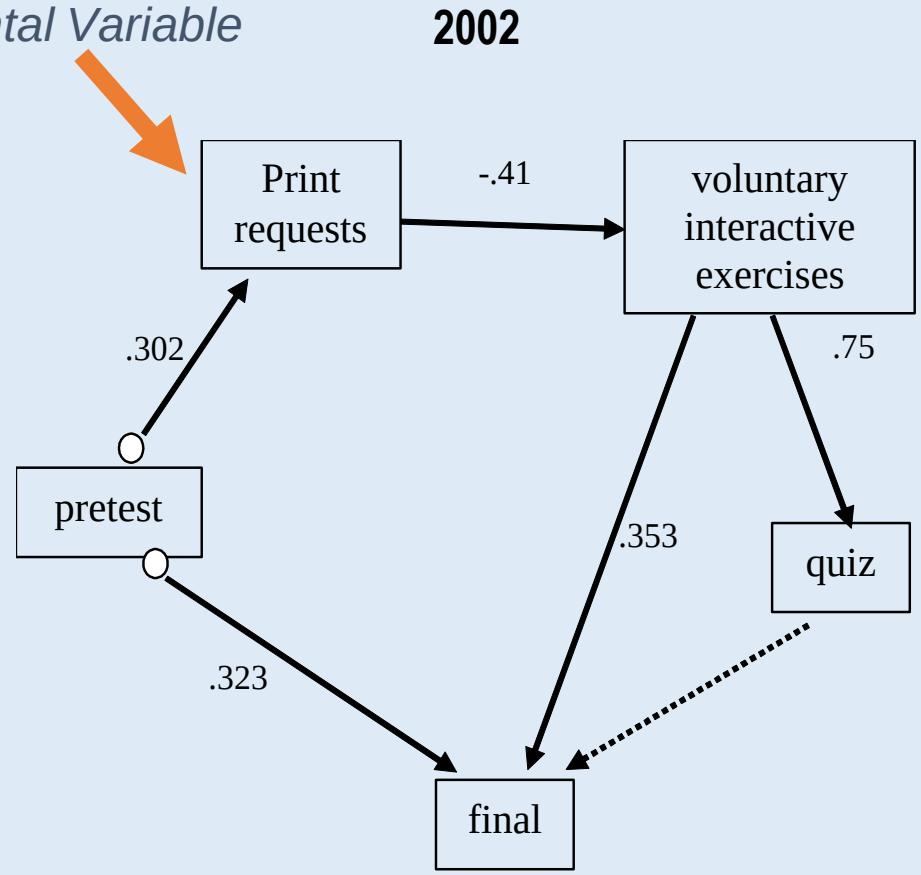


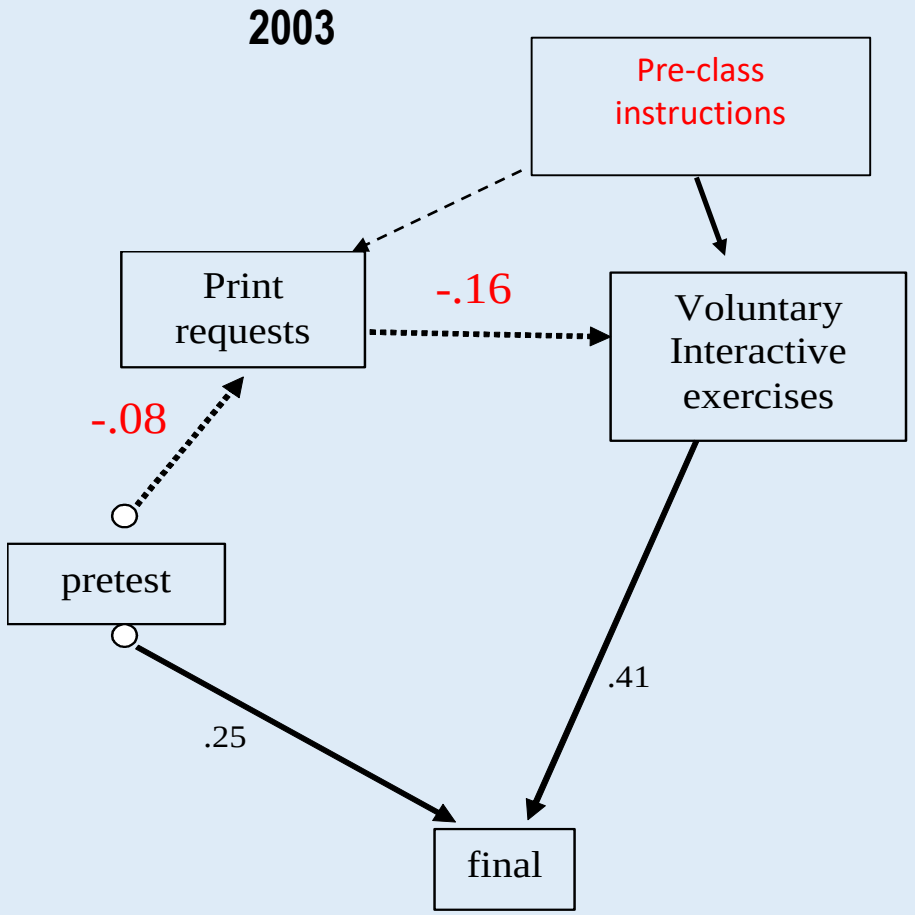
15 Variables

- Pre-test (%)
- Print-outs (% modules printed)
- Quiz Scores (avg. %)
- Voluntary Exercises (% completed)
- Final Exam (%)
- 9 other variables

Voluntary interaction → Learning: "The Doer Effect" 2005

Detectable Instrumental Variable



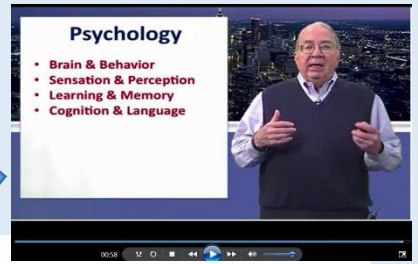


Scheines, R., Leinhardt, G., Smith, J., and Cho, K. (2005) "Replacing Lecture with Web-Based Course Materials, *Journal of Educational Computing Research*, 32, 1, 1-26.

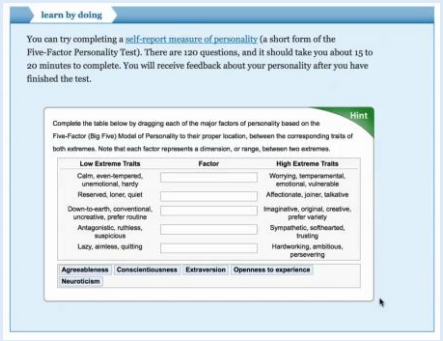
"The Doer Effect" 2015

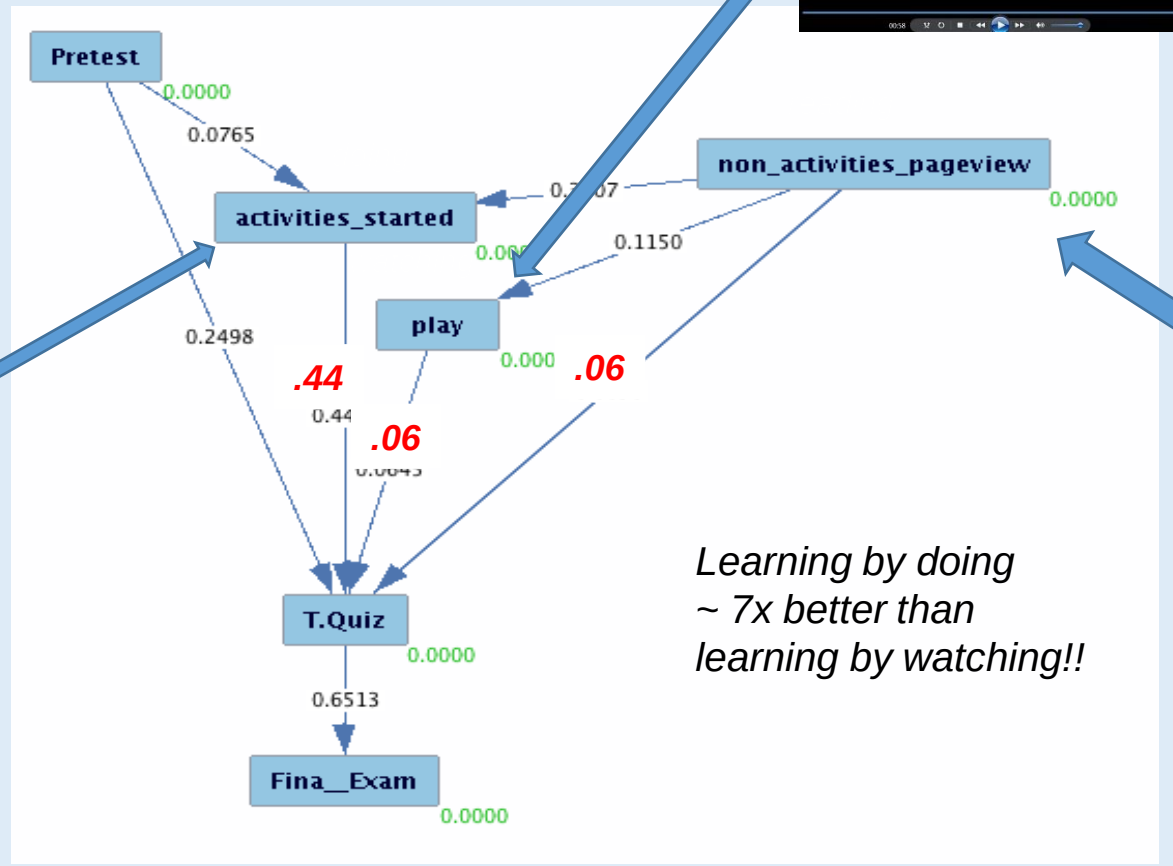
What student choices in an online course cause the most learning?

Watching lecture video

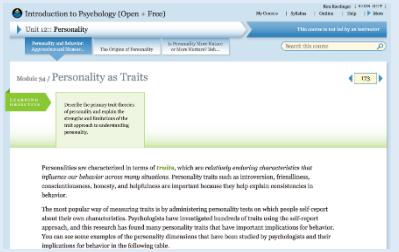


Doing online activities with hints & feedback



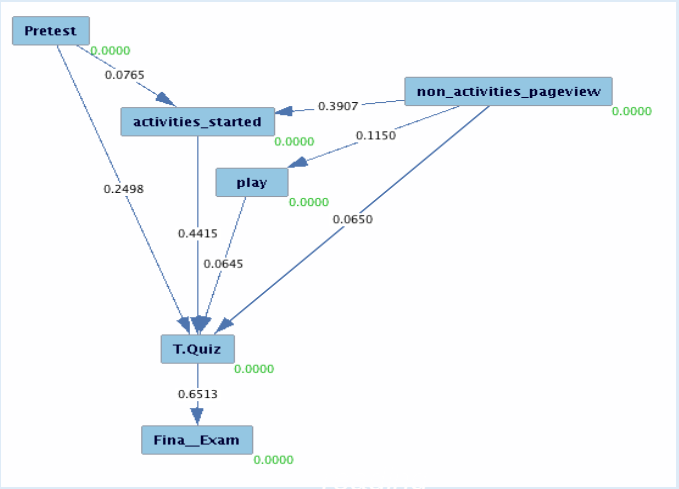


Reading web pages



Learning by doing ~ 7x better than learning by watching!!

Koedinger, Kim, Jia, McLaughlin, & Bier (2015). Learning is not a spectator sport: Doing is better than watching for learning from a MOOC. In *Proceedings of the Second ACM Conference on Learning at Scale*.

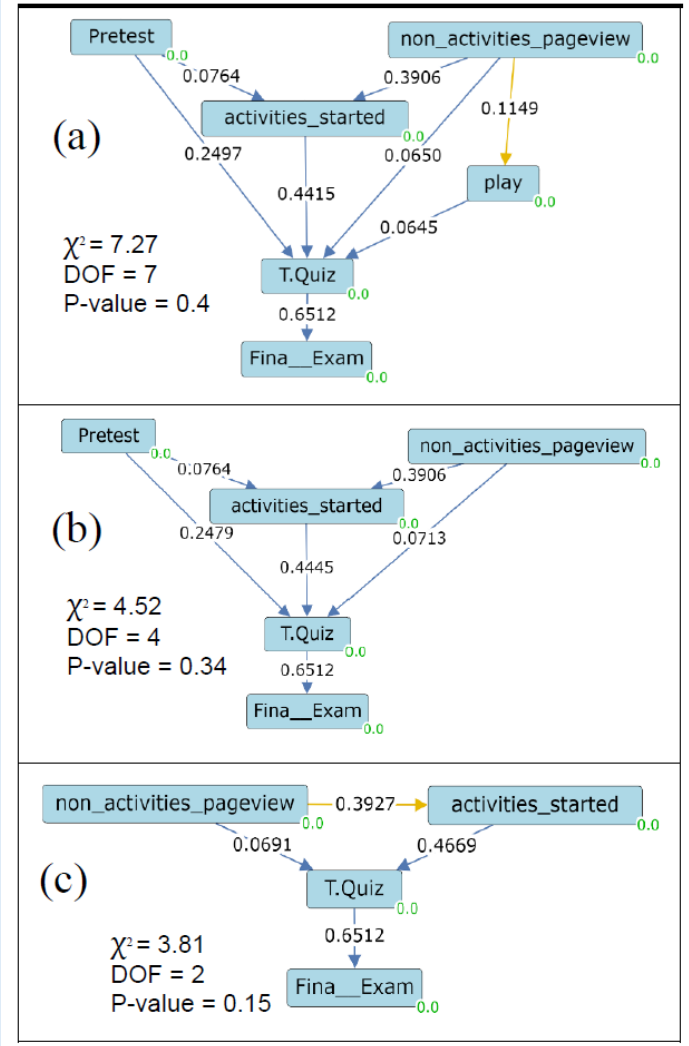


"The Doer Effect" 2018

Table 2. The causal model that was discovered on GTech's psychology dataset was estimated using data from datasets listed in the first row of the table.

	UMUC Biology	UMUC Info Sci	UMUC Psychology	UMUC Statistics	C@CM	UMUC Biology (sample)	UMUC Info Sci (sample)
#Students	3516	6112	89	61	383	300	300
Chi-square	78.89	18.44	1.04	28.33	14.30	11.92*	3.32*
DOF	2	2	2	2	4	2	2
P-value	0	0	0.59	0	0	0.02*	0.49*

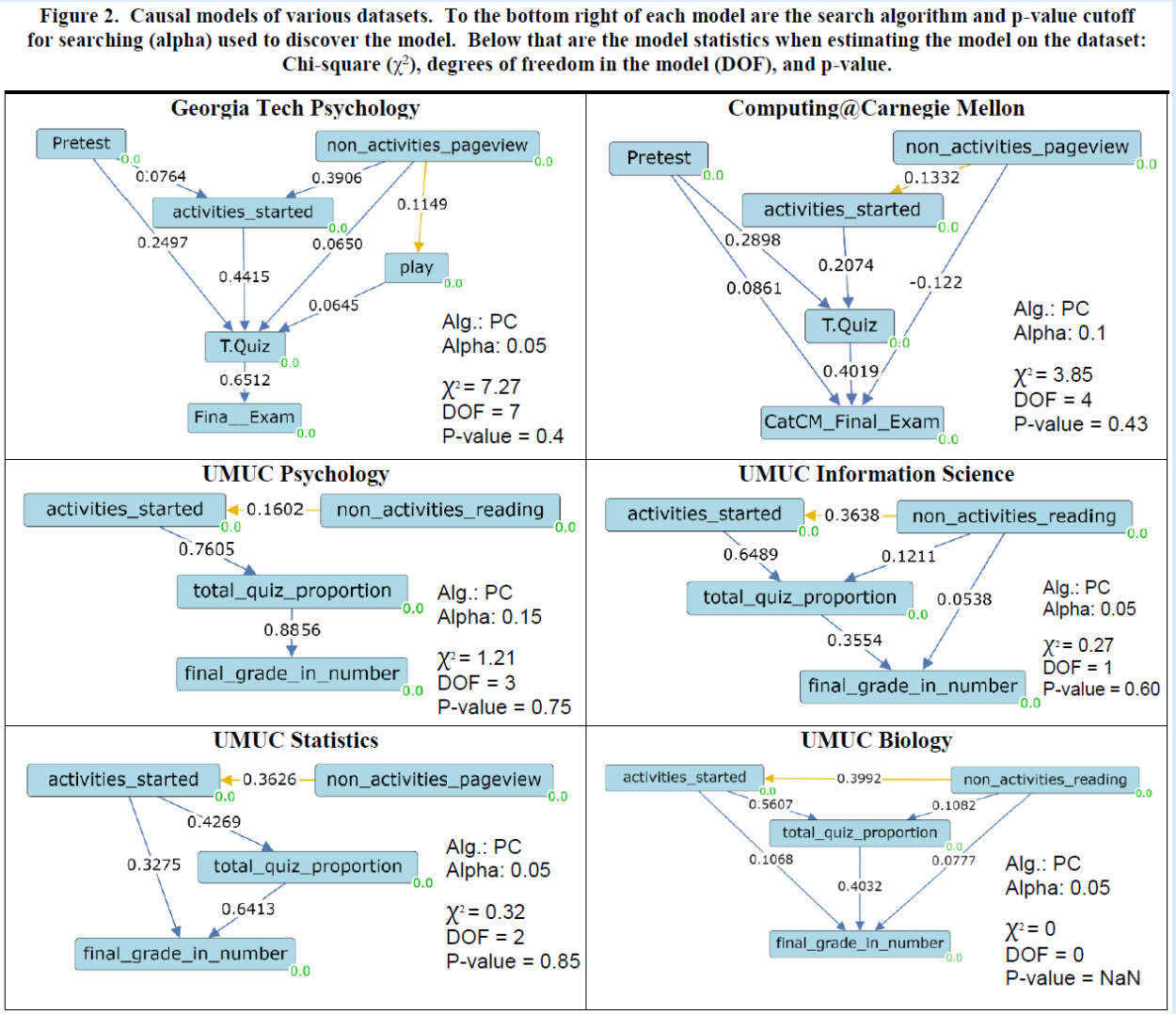
Figure 1. Using subsets of variables from the Georgia Tech Psychology dataset, three causal models were discovered using the PC algorithm and an alpha value of 0.05 as in [6].



Koedinger, K., Scheines, R., and Schaldebrand, P. (2018). "Is the Doer Effect Robust Across Multiple Data Sets?", *Proceedings of the 11th Intl. Conference on Educational Data Mining*, Buffalo, NY.

Is the Doer Effect Robust Across Multiple Data Sets?

Figure 2. Causal models of various datasets. To the bottom right of each model are the search algorithm and p-value cutoff for searching (alpha) used to discover the model. Below that are the model statistics when estimating the model on the dataset: Chi-square (χ^2), degrees of freedom in the model (DOF), and p-value.



Yes – in all samples doing was caually connected to success, and much more strongly than reading or watching so

Subsequent experiments confirm this finding – and indicate it is underestimated by 2x (unpublished)

Koedinger, K., Scheines, R., and Schaldebrand, P. (2018). "Is the Doer Effect Robust Across Multiple Data Sets?", *Proceedings of the 11th Intl. Conference on Educational Data Mining*, Buffalo, NY.

Fractions Tutor (5th grade)

Naming Fractions

A Let's name a fraction to compare it to another!

This is the unit of the fraction shown below. Let's find the fraction of the blue sections.

1 

How many blue sections are there? 3

2 Use the arrows to make all sections the same size.

3 How many blue sections do you need to fill the gray diagram? 8

4 What fraction are the blue sections of the unit? $\frac{3}{8}$

B Let's name a second fraction to compare it to the first!

This is the unit of the fraction shown below. Let's find the fraction of the purple sections.

1 

How many purple sections are there? 3

2 Use the arrows to make all sections the same size.

3 How many purple sections do you need to fill the gray diagram? 4

4 What fraction are the purple sections of the unit? $\frac{3}{4}$

C Are these fractions still the same?

1 How many total sections are in the blue diagram? 8
How many total sections are in the purple diagram? 4

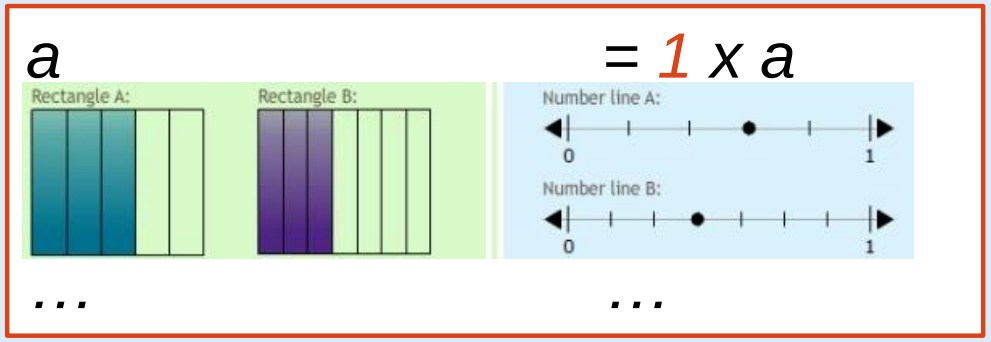
2 The blue sections are equal to the purple section.

3 However, the diagrams show different fractions, because their units have different shapes.

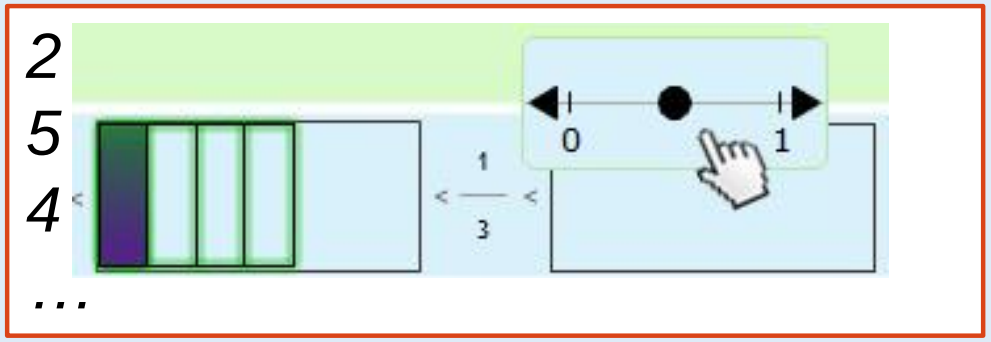
? Hint

Fluency vs. Understanding

- Understanding: sense-making processes

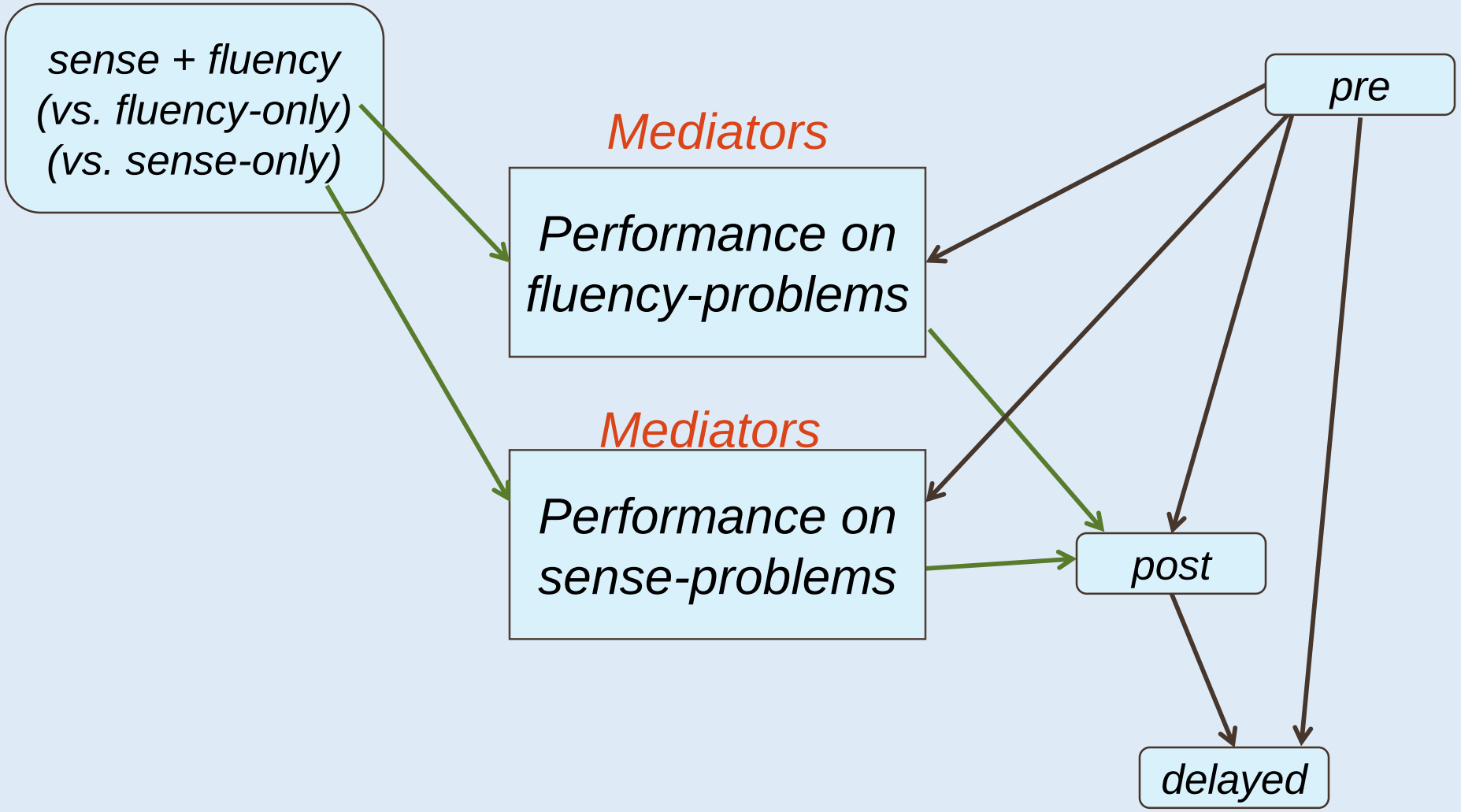


- Fluency: fast, compiled, reliable



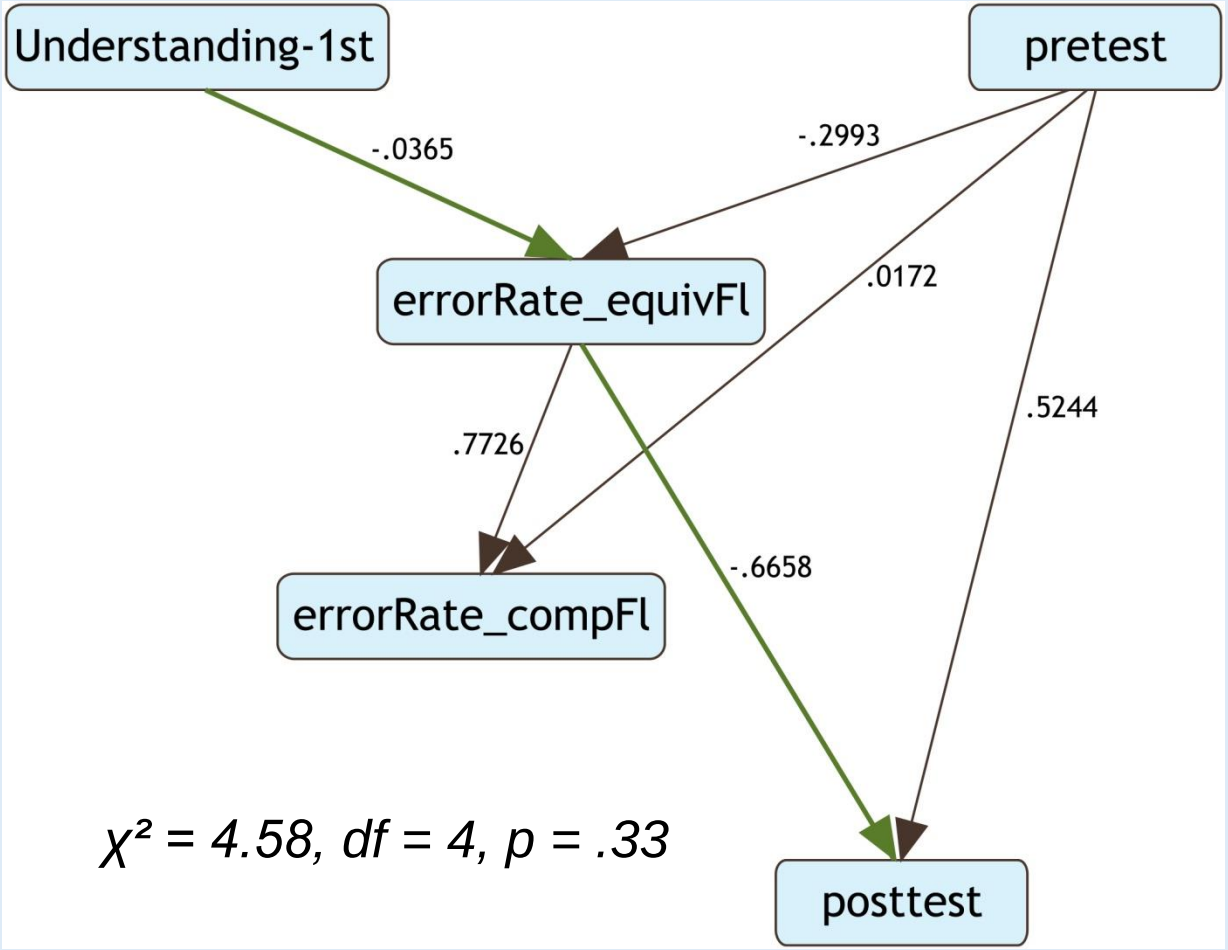
Which to teach first?

Mediation Hypotheses



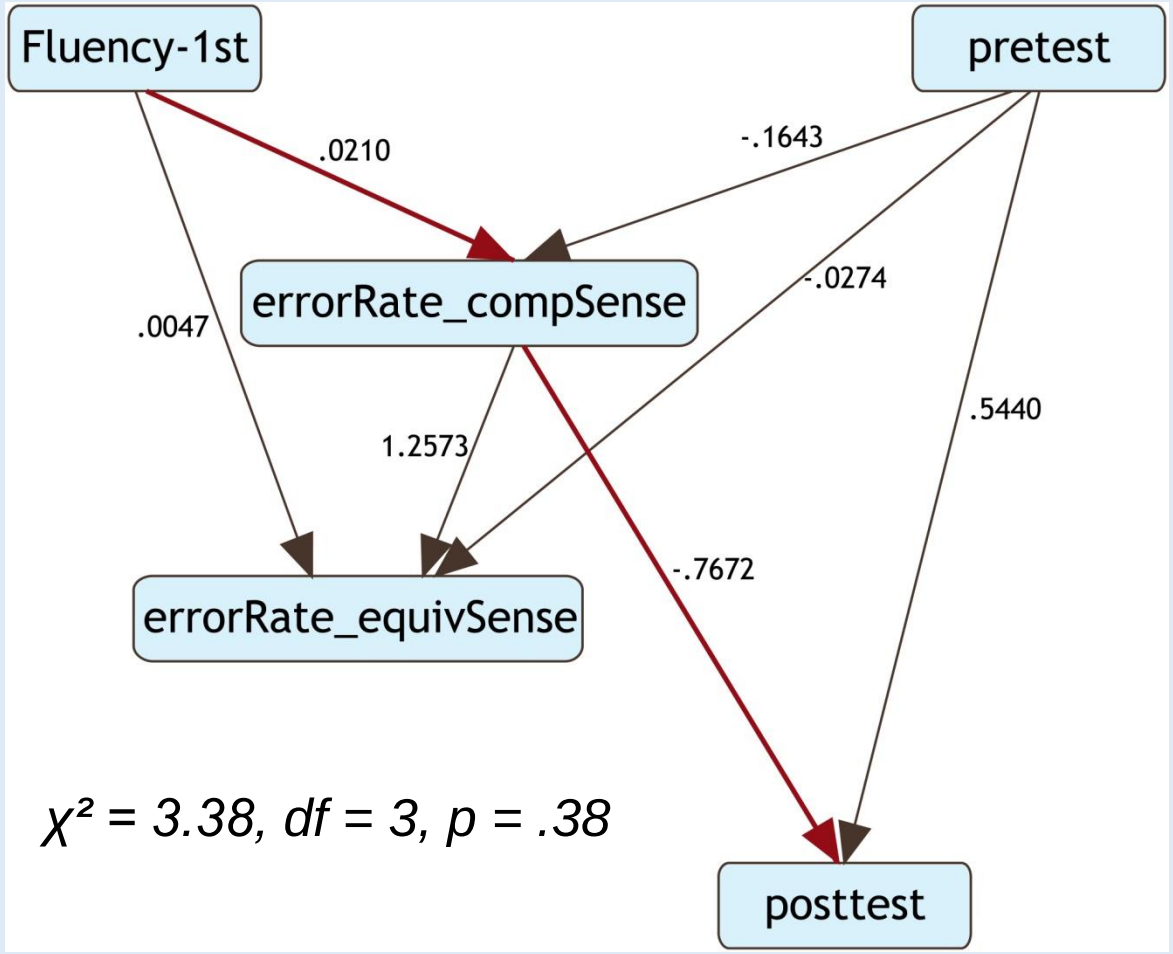
Model Search Results

Understanding-1st *reduces* Fluency errors – which in turn *increases* post-test



Model Search Results

Fluency-1st *increases* Understanding errors – which in turn *decreases* post-test



Educational Research

Follow up Experiment: Sense-making first vs. Fluency-first

Sense-making first dramatically outperformed fluency-first condition on fraction learning

Educational Data → Causal Discovery Methods
→ Instrumental Variable Detection
→ Causes of Educational Returns

Discovering Causal Models with Optimization: Confounders, Cycles, and Instrument Validity

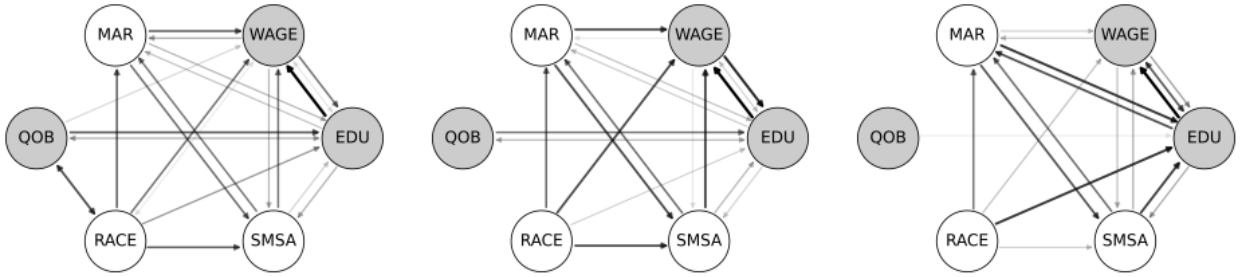
Frederick Eberhardt
Division of Humanities and Social Sciences, California Institute of Technology, fde@caltech.edu.

Nur Kaynar*
Samuel Curtis Johnson Graduate School of Management, Cornell University, sk2739@cornell.edu.

Auyon Siddiq
Anderson School of Management, University of California, Los Angeles, auyon.siddiq@anderson.ucla.edu.

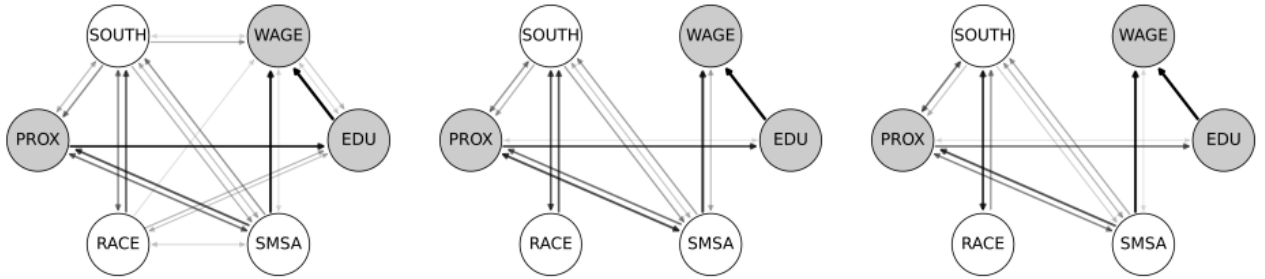
We propose a new optimization-based method for learning causal structures from observational data, a process known as *causal discovery*. Our method takes as input observational data over a set of variables and returns a graph in which causal relations are specified by directed edges. We consider a highly general search space that accommodates latent confounders and feedback cycles, which few extant methods do. We formulate the discovery problem as an integer program, and propose a solution technique that exploits the conditional independence structure in the data to identify promising edges for inclusion in the output graph. In the large-sample limit, our method recovers a graph that is (Markov) equivalent to the true data-generating graph. Computationally, our method is competitive with the state-of-the-art, and can solve in minutes instances that are intractable for alternative causal discovery methods. We leverage our method to develop a graphical test for the validity of an instrumental variable and demonstrate it on the influential instruments for estimating the returns to education from Angrist and Krueger (1991) and Card (1993). In particular, our test complements existing instrument tests by revealing the precise causal pathways that undermine instrument validity, highlighting the unique merits of the graphical perspective on causality.

Educational Data → Causal Discovery Methods → Instrumental Variable Detection → Causes of Educational Returns



(a) $c_{0.95}$ (b) c_1 (c) $c_{1.05}$

Figure 6 Edge frequencies over 50 bootstrap repetitions of EdgeGen applied to Angrist and Krueger (1991) data for three levels of complexity penalty. Only edges with frequency ≥ 0.1 are shown.



(a) $c_{0.95}$ (b) c_1 (c) $c_{1.05}$

Figure 7 Edge frequencies over 50 bootstrap repetitions of EdgeGen applied to Card (1993) data for three levels of complexity penalty. Only edges with frequency ≥ 0.1 are shown.

Causal Discovery with Models: Behavior, Affect, and Learning in Cognitive Tutor Algebra

Stephen E. Fancsali
Carnegie Learning, Inc.

Proceedings of the 7th International Conference on Educational Data Mining (EDM 2014)

ABSTRACT

Non-cognitive and behavioral phenomena, including gaming the system, off-task behavior, and affect, have proven to be important for understanding student learning outcomes. The nature of these phenomena requires investigations into their causal structure. For example, given that gaming the system has been associated with poorer learning outcomes, would reducing such behavior improve outcomes? Answering this question requires an understanding of whether gaming the system is a cause of poor outcomes, rather than, for example, only sharing a common cause with factors influencing learning. Because controlled experiments to settle such causal questions are often costly or impractical, we employ algorithmic search for the structure of graphical causal models from non-experimental data. Using sensor-free, data-driven detectors of behavior and affect, this work extends Baker and Yacef's notion of "discovery with models" to incorporate causal discovery and reasoning, resulting in an approach we call "causal discovery with models." We explore a case study of this approach using data from Carnegie Learning's Cognitive Tutor for Algebra and raise questions for future research.

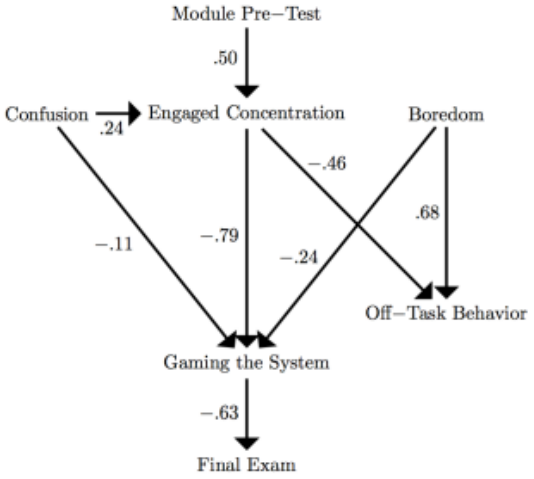


Figure 2. Estimated linear structural equation model

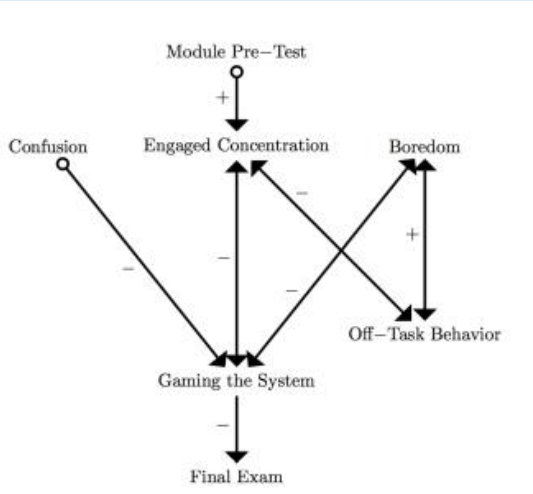


Figure 3. PAG causal model learned with weaker assumptions

Brain Research

fMRI → Causal Discovery Methods → Processing Mechanisms

NeuroImage 49 (2010) 1545–1558

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Six problems for causal inference from fMRI

J.D. Ramsey [a,*], S.J. Hanson [b], C. Hanson [b], Y.O. Halchenko [b], R.A. Poldrack [c], C. Glymour [d]

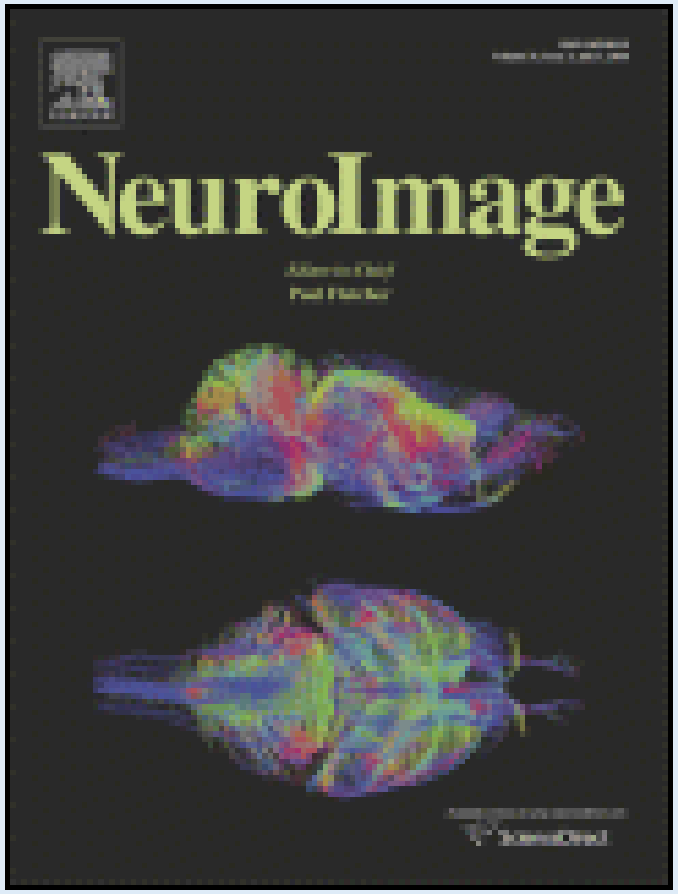
[a] Department of Philosophy, Carnegie Mellon University, Pittsburgh, PA 15213
[b] Department of Psychology, Rutgers University, Rumba Lab
[c] Imaging Research Center and Departments of Psychology and Neurobiology, University of Texas at Austin
[d] Department of Philosophy, Carnegie Mellon University, and Florida Institute for Human and Machine Cognition

ARTICLE INFO

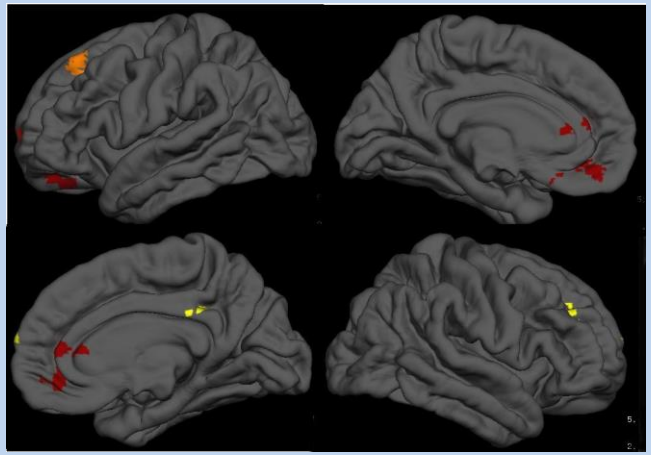
Article history:
Received 13 February 2009
Revised 7 August 2009
Accepted 31 August 2009
Available online 9 September 2009

ABSTRACT

Neuroimaging (e.g. fMRI) data are increasingly used to attempt to identify not only brain regions of interest (ROIs) that are especially active during perception, cognition, and action, but also the qualitative causal relations among activity in these regions (known as *effective connectivity*; Friston, 1994). Previous investigations and anatomical and physiological knowledge may somewhat constrain the possible hypotheses, but there often remains a vast space of possible causal structures. To find actual effective connectivity relations, search methods must accommodate indirect measurements of nonlinear time series dependencies, feedback, multiple subjects possibly varying in identified regions of interest, and unknown possible location-dependent variations in BOLD response delays. We describe combinations of procedures that under these conditions find feed-forward sub-structure characteristic of a group of subjects. The method is illustrated with an empirical data set and confirmed with simulations of time series of non-linear, randomly generated, effective connectivities, with feedback, subject to random differences of BOLD delays, with regions of interest missing at random for some subjects, measured with noise approximating the signal to noise ratio of the empirical data.

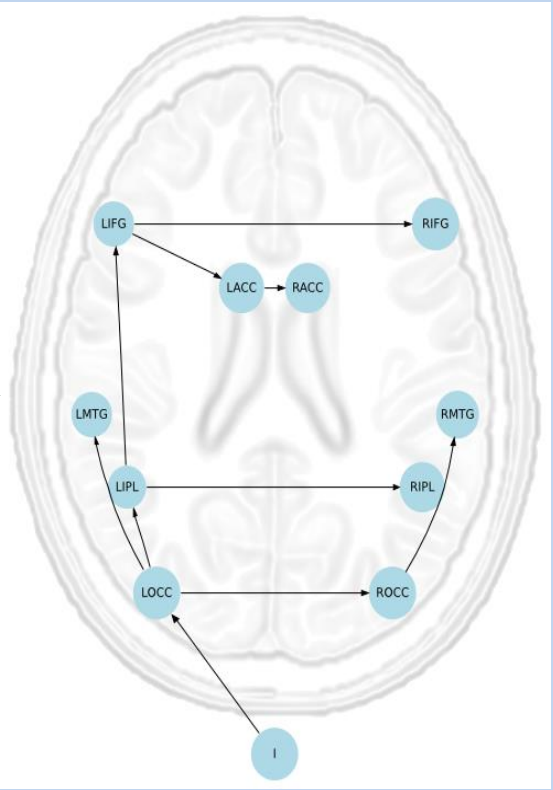


fMRI
(~44,000 voxels)



(ROI)
~10-20 Regions of Interest

Causal network discovery

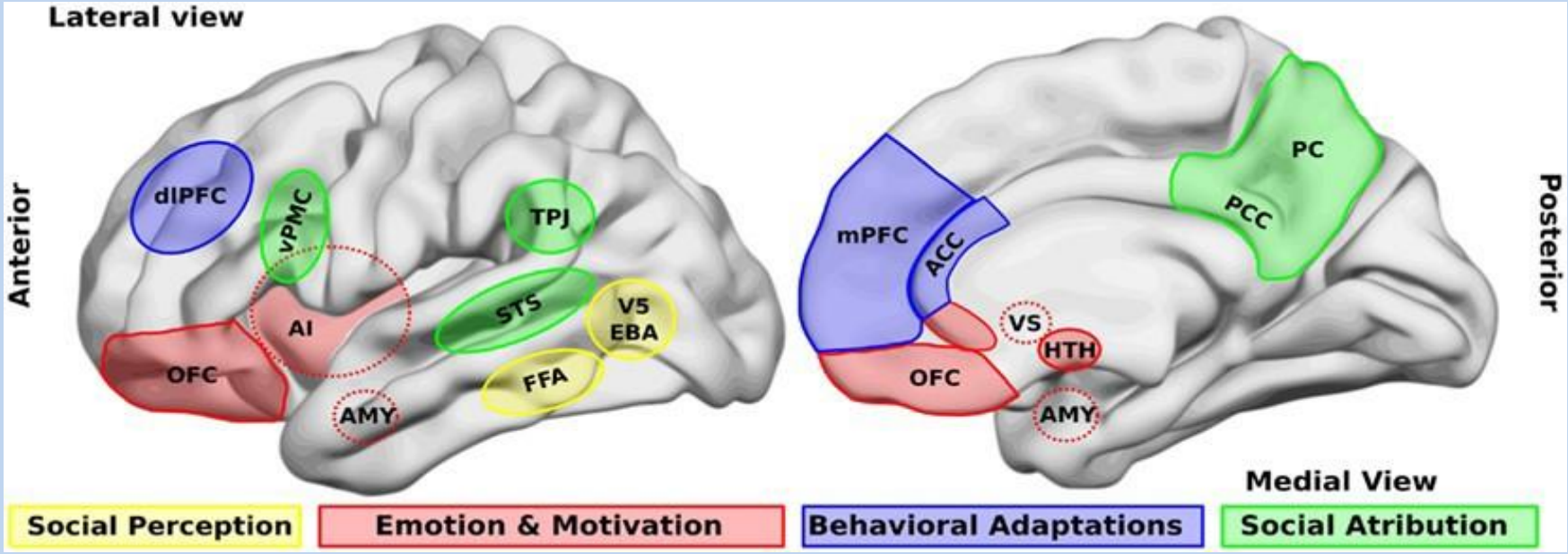


Autism

Catherine Hanson, Rutgers

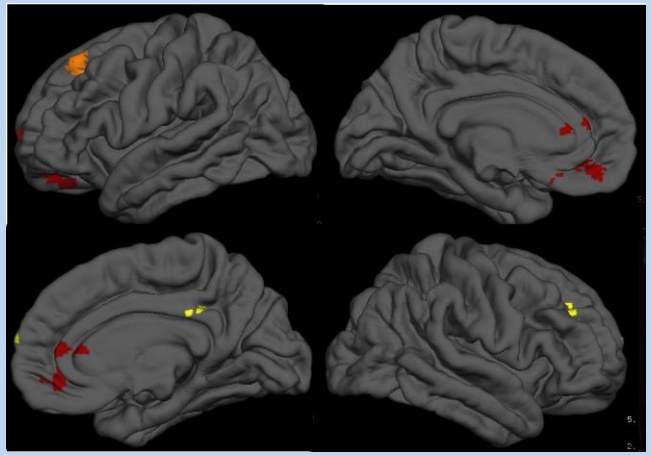
Subjects: Autistic Spectrum Disorder vs. Neurotypical

Usual Approach:
Search for differential recruitment of brain regions



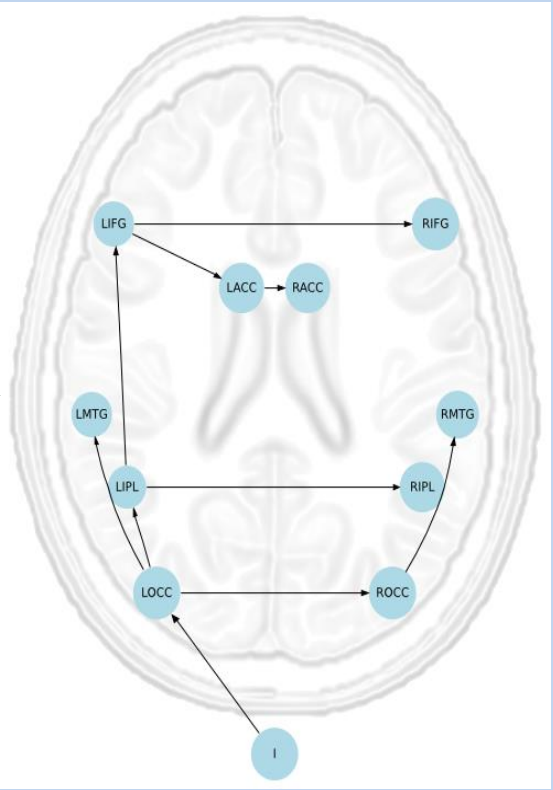
fMRI
(~44,000 voxels)

- Face processing task
- Theory of Mind task
- Action understanding task



(ROI)
~10-20 Regions of Interest

Causal network discovery

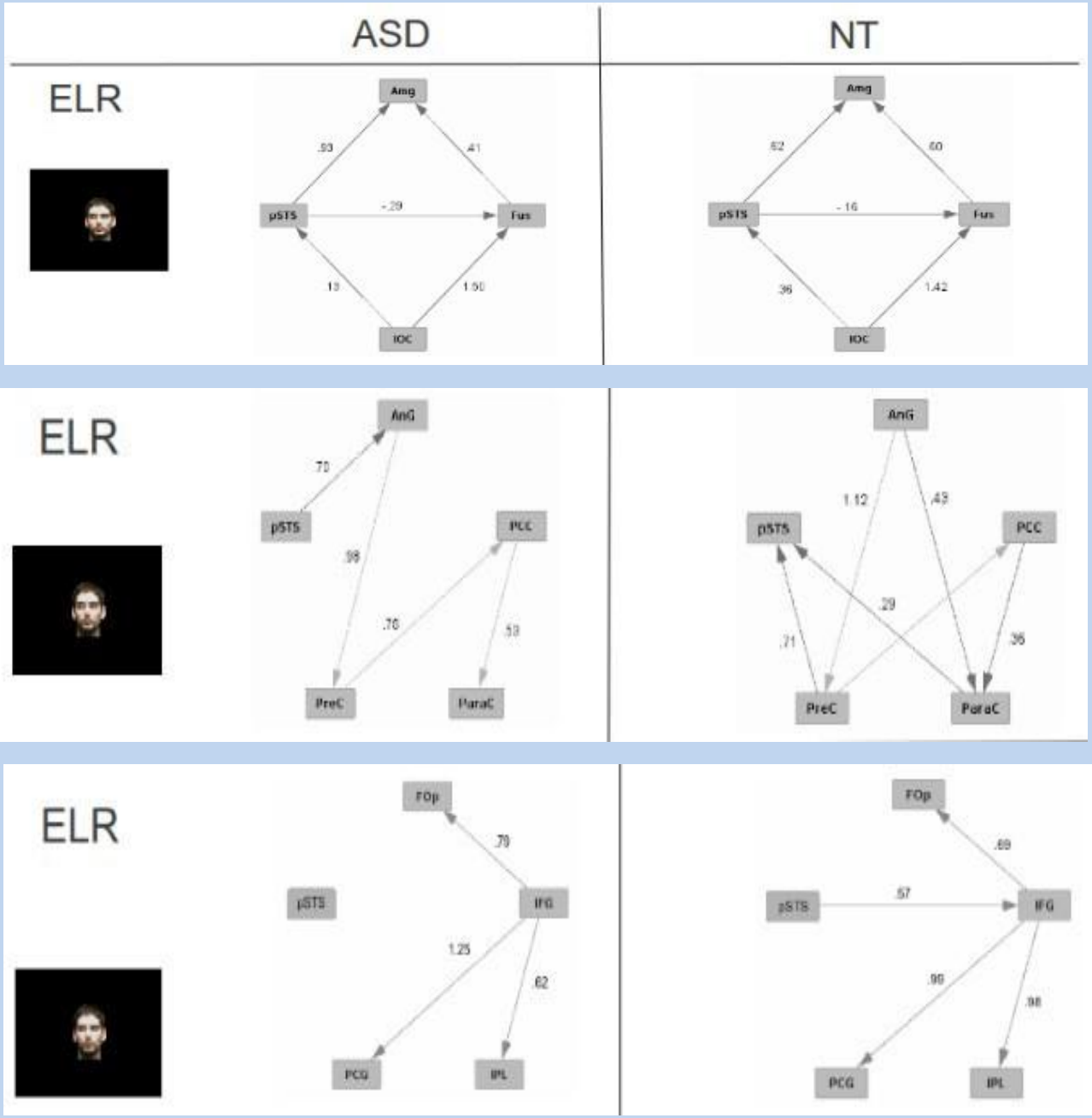


Results

FACE recognition

Theory of Mind

ACTION



Causal Discovery → Autism

Biwei Huang, CMU & UCSD

ACADEMIC PRESS

Neural Engineering Techniques for Autism Spectrum Disorder

Volume 1: Imaging and Signal Analysis

2021, Pages 237-267

Chapter 12 - Diagnosis of autism spectrum disorder by causal influence strength learned from resting-state fMRI data

Biwei Huang

Similar Follow on Work Catherine Hanson, Rutgers

Published in final edited form as:
J Alcohol Drug Depend. 2017 August ; 5(4): . doi:10.4172/2329-6488.1000279.

Modeling Causal Relationships among Brain Areas in the Mesocorticolimbic System during Resting-State in Cocaine Users Utilizing a Graph Theoretic Approach

Suchismita Ray[1,*], **Bharat B Biswal**[2], **Ashley Aya**[1], **Suril Gohel**[3], **Aradhana Srinagesh**[1], **Catherine Hanson**[4], and **Stephen J. Hanson**[4]

Results—The causal interaction pattern was different between the two groups. The feed-forward pattern found in cocaine smokers, between 7 ROIs of the MCLS during resting-state [ventral tegmental area (VTA)→hippocampus (HIPP)→ventral striatum (VenStri)→orbital frontal cortex (OFC), medial frontal cortex (MFC), anterior cingulate cortex (ACC), dorsolateral prefrontal cortex (DLPFC)], was absent in controls. That is, the subcortical VenStri area had a causal influence on four cortical brain areas only in cocaine users.

ELSEVIER

Drug and Alcohol Dependence
Volume 146, 1 January 2015, Page e77

Modeling causal relationship between memory and craving-related brain networks in non-treatment seeking cocaine smokers using images, a graph theoretic approach

Suchismita Ray [1], Catherine Hanson [2], Margaret Haney [3], Bharat Biswal [4], Stephen J. Hanson [2 5]

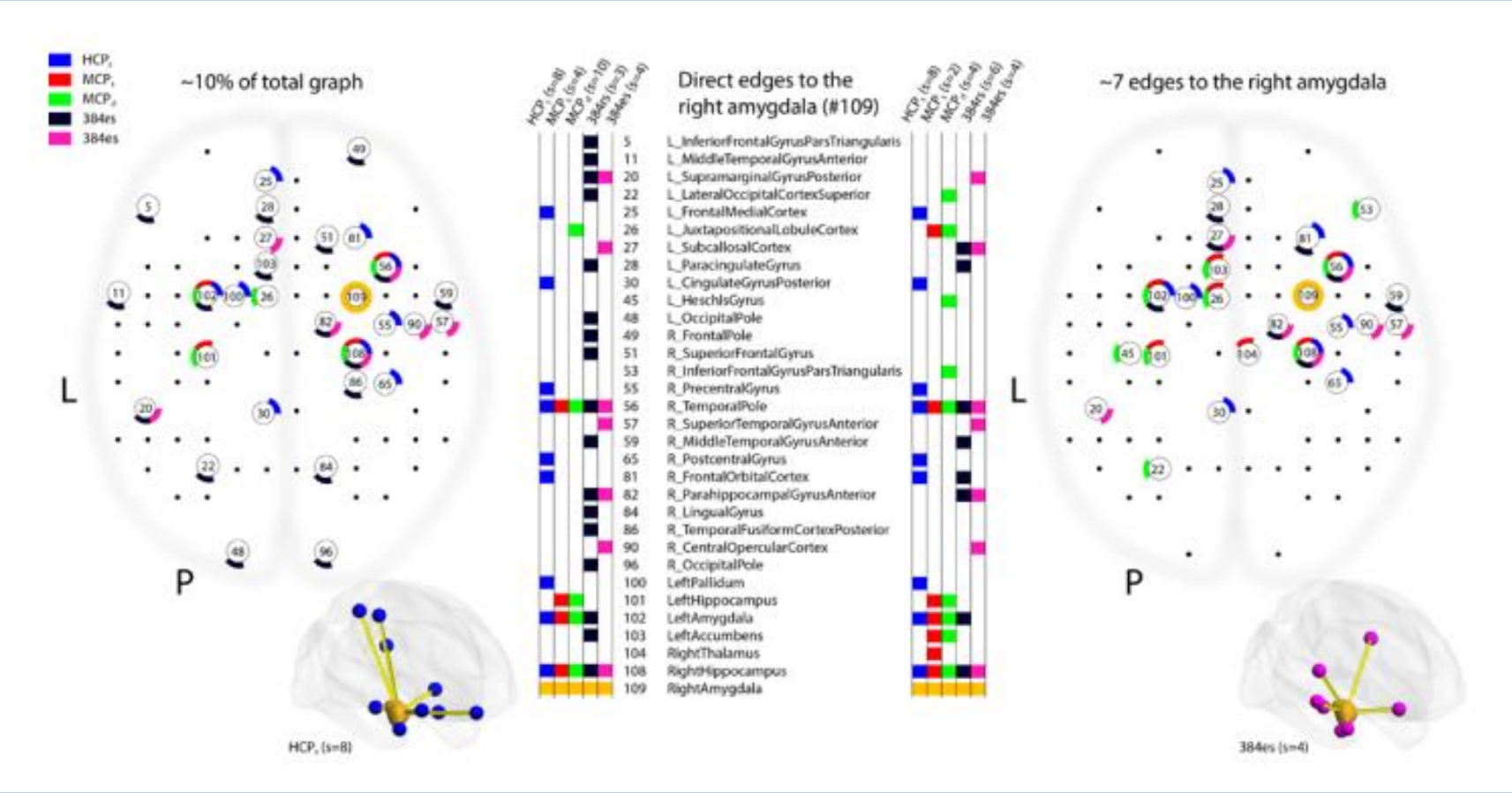
es-fMRI Data → Causal Discovery (FGES) → Emotion Network Structure (Amygdala)

Causal Mapping of Emotion Networks in the Human Brain: Framework and Preliminary Findings

Julien Dubois[1,4], Hiroyuki Oya[5], J. Michael Tyszka[2], Matthew Howard III[5], Frederick Eberhardt[1], and Ralph Adolphs[1,2,3]

[1]Division of Humanities and Social Sciences,
[2]Division of Biology, and
[3]Chen Neuroscience Institute, California Institute of Technology, Pasadena CA 91125, USA
[4]Department of Neurosurgery, Cedars-Sinai Medical Center, Los Angeles, CA 90048, USA
[5]Department of Neurosurgery, Human Brain Research Laboratory, University of Iowa, IA 52241, USA

es-fMRI Data → Causal Discovery (FGES) → Emotion Network Structure (Amygdala)



Climate Research

Climate Teleconnections

Journal of Machine Learning Research 9 (2008) 967-991 Submitted 9/06; Revised 9/07; Published 5/08

Search for Additive Nonlinear Time Series Causal Models

Tianjiao Chu TIC19@PITT.EDU
Department of Obstetrics, Gynecology & Reproductive Sciences
University of Pittsburgh
204 Craft Ave., Room B409
Pittsburgh, PA 15213, USA

Clark Glymour CG09@ANDREW.CMU.EDU
Department of Philosophy
Carnegie Mellon University
Pittsburgh, PA 15213, USA

Editor: Greg Ridgeway

Abstract

Pointwise consistent, feasible procedures for estimating contemporaneous linear causal structure from time series data have been developed using multiple conditional independence tests, but no such procedures are available for non-linear systems. We describe a feasible procedure for learning a class of non-linear time series structures, which we call additive non-linear time series. We show that for data generated from stationary models of this type, two classes of conditional independence relations among time series variables and their lags can be tested efficiently and consistently using tests based on additive model regression. Combining results of statistical tests for these two classes of conditional independence relations and the temporal structure of time series data, a new consistent model specification procedure is able to extract relatively detailed causal information. We investigate the finite sample behavior of the procedure through simulation, and illustrate the application of this method through analysis of the possible causal connections among four ocean indices. Several variants of the procedure are also discussed.

Keywords: conditional independence test, contemporaneous causation, additive model regression, Granger causality, ocean indices

Climate Teleconnections

Our data set consists of the following 4 ocean climate indices, recorded monthly from 1958 to 1999, each forming a time series of 504 time steps:

SOI Southern Oscillation Index: Sea Level Pressure (SLP) anomalies between Darwin and Tahiti

WP Western Pacific: Low frequency temporal function of the 'zonal dipole' SLP spatial pattern over the North Pacific.

AO Arctic Oscillation: First principal component of SLP poleward of 20° N

NAO North Atlantic Oscillation: Normalized SLP differences between Ponta Delgada, Azores and Stykkisholmur, Iceland

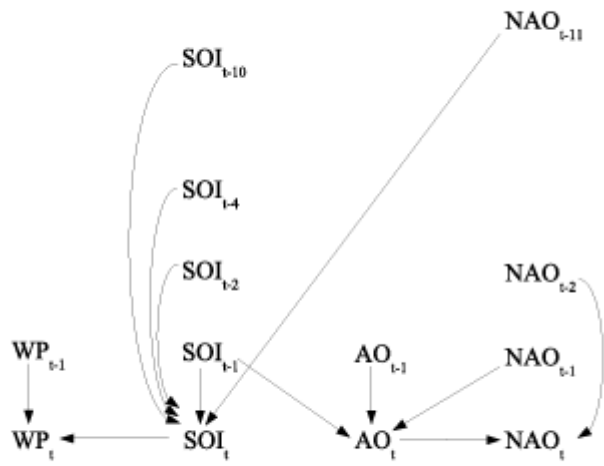


Figure 7: Causal connections among 4 ocean climate indices, using the additive non-linear algorithm

Climate Research

Time Series Climate Data → Causal Discovery Methods
→ Causal Mechanisms in Earth's Climate

nature communications

Explore content ⌄ About the journal ⌄ Publish with us ⌄

nature › nature communications › perspectives › article

Perspective | Open Access | Published: 14 June 2019

Inferring causation from time series in Earth system sciences

Jakob Runge ✉, Sebastian Bathiany, Erik Bollt, Gustau Camps-Valls, Dim Coumou, Ethan Deyle, Clark Glymour, Marlene Kretschmer, Miguel D. Mahecha, Jordi Muñoz-Marí, Egbert H. van Nes, Jonas Peters, Rick Quax, Markus Reichstein, Marten Scheffer, Bernhard Schölkopf, Peter Spirtes, George Sugihara, Jie Sun, Kun Zhang & Jakob Zscheischler

Nature Communications **10**, Article number: 2553 (2019) | Cite this article

Time Series Climate Data → Causal Discovery Methods → Causal Structure of the Climate

Jakob Runge

FOLLOW

German Aerospace Center and TU Berlin
Verified email at dlr.de - Homepage

Causal Inference Time Series Statistical Machine Learning Information Theory Earth Science

TITLE	CITED BY	YEAR
Inferring causation from time series in Earth system sciences J Runge, S Bathiany, E Bollt, G Camps-Valls, D Coumou, E Deyle, ... Nature communications 10 (1), 2553	501	2019
Detecting and quantifying causal associations in large nonlinear time series datasets J Runge, P Nowack, M Kretschmer, S Flaxman, D Sejdinovic Science advances 5 (11), eaau4996	481	2019
Escaping the curse of dimensionality in estimating multivariate transfer entropy J Runge, J Heitzig, V Petoukhov, J Kurths Physical review letters 108 (25), 258701	314	2012
Causal network reconstruction from time series: From theoretical assumptions to practical estimation J Runge Chaos: An Interdisciplinary Journal of Nonlinear Science 28 (7)	269	2018
Identifying causal gateways and mediators in complex spatio-temporal systems J Runge, V Petoukhov, JF Donges, J Hlinka, N Jajcay, M Vejmelka, ... Nature communications 6 (1), 8502	246	2015
Using causal effect networks to analyze different Arctic drivers of midlatitude winter circulation M Kretschmer, D Coumou, JF Donges, J Runge Journal of climate 29 (11), 4069-4081	235	2016
Quantifying the strength and delay of climatic interactions: The ambiguities of cross correlation and a novel measure based on graphical models J Runge, V Petoukhov, J Kurths Journal of climate 27 (2), 720-739	172	2014

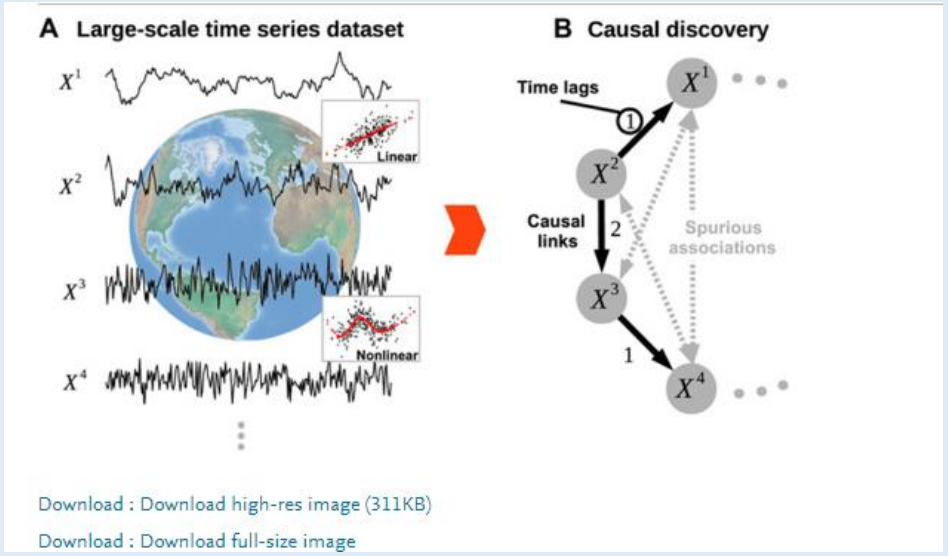
Time Series Climate Data → Causal Discovery Methods → Causal Structure of the Climate

Jakob Runge FOLLOW

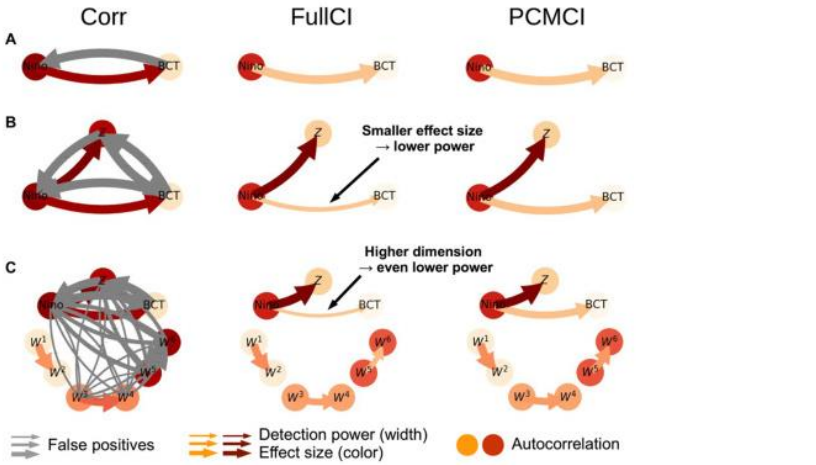
German Aerospace Center and TU Berlin
Verified email at dlr.de - Homepage

Causal Inference Time Series Statistical Machine Learning Information Theory Earth Science

TITLE	CITED BY	YEAR
Inferring causation from time series in Earth system sciences J Runge, S Bathiany, E Bollt, G Camps-Valls, D Coumou, E Deyle, ... Nature communications 10 (1), 2553	501	2019
Detecting and quantifying causal associations in large nonlinear time series datasets J Runge, P Nowack, M Kretschmer, S Flaxman, D Sejdinovic Science advances 5 (11), eaau4996	481	2019



Download : Download high-res image (311KB)
Download : Download full-size image



Download : Download high-res image (434KB)
Download : Download full-size image

Fig. 2. Source: Fig.2 ofRunge et al. (2019) on the higher detection power of PCMCI partial correlation compared to Correlation, and FullCI partial correlation.

ELSEVIER

Energy Reports

Volume 7, November 2021, Pages 6196-6204

Research paper

Exploring nonlinearity on the CO_2 emissions, economic production and energy use nexus: A causal discovery approach

Peter Martey Addo [a], Christelle Manibialoa [a], Florent McIsaac [b]

Show more

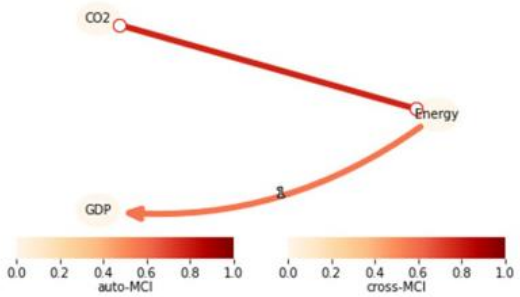
+ Add to Mendeley Share Cite

https://doi.org/10.1016/j.egyr.2021.09.026

Get rights and content

Highlights

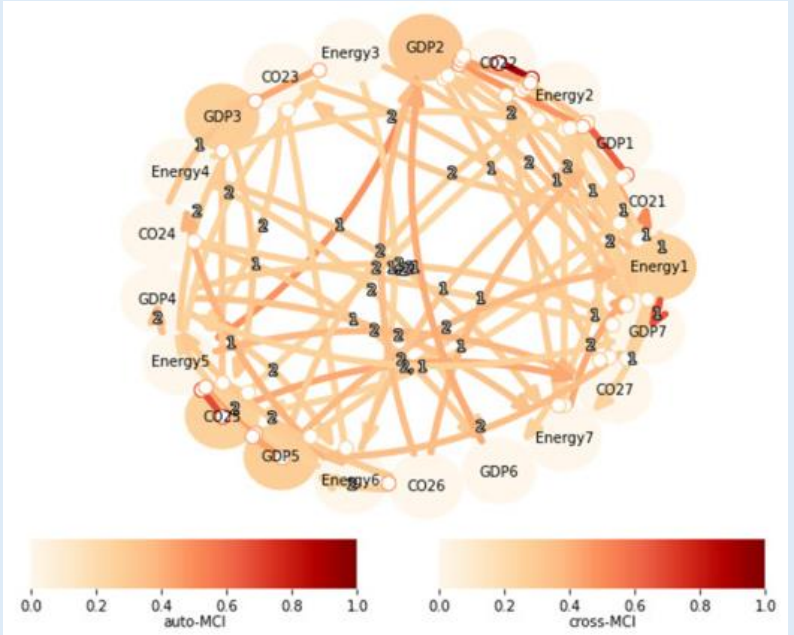
- At the global scale, energy is central in the economic and decarbonization debate.
- Coordination between countries will improve policy effectiveness on climate actions.
- Energy and climate policies in three regions will impact global economic dynamics.
- Examining interactions between carbon emissions, economic production, and energy use.
- Too rapid a transition to net-zero emissions in the energy sector may have negative consequences for economic growth.



Download : Download high-res image (93KB)

Download : Download full-size image

Fig. 5. Causal dependence at global level. The PCMCI with CMIknn independence test produces better predictive power with $NRMSE = 0.77$ compared to Parcorr and GPDC.



Download : Download high-res image (477KB)

Download : Download full-size image

Fig. 7. Causal dependence of all regions. The PCMCI with GPDC indepe

Journal of the Atmospheric Sciences

A Causality-Based View of the Interaction between Synoptic- and Planetary-Scale Atmospheric Disturbances

Savini M. Samarasinghe, Yi Deng, and Imme Ebert-Uphoff

Online Publication: 01 Feb 2020

Print Publication: 01 Mar 2020

DOI: https://doi.org/10.1175/JAS-D-18-0163.1

Page(s): 925–941

Article History | Download PDF | © Get Permissions

Abstract/Excerpt | Full Text | PDF

Abstract

This paper reports preliminary yet encouraging findings on the use of causal discovery methods to understand the interaction between atmospheric planetary- and synoptic-scale disturbances in the Northern Hemisphere. Specifically, constraint-based structure learning of probabilistic graphical models is applied to the spherical harmonics decomposition of the daily 500-hPa geopotential height field in boreal winter for the period 1948–2015. Active causal pathways among different spherical harmonics components are identified and documented in the form of a temporal probabilistic graphical model. Since, by definition, the structure learning algorithm used here only robustly identifies linear causal effects, we report only causal pathways between two groups of disturbances with sufficiently large differences in temporal and/or spatial scales, that is, planetary-scale (mainly zonal wavenumbers 1–3) and synoptic-scale disturbances (mainly zonal wavenumbers 6–8). Daily reconstruction of geopotential heights using only interacting scales suggest that the modulation of synoptic-scale disturbances by planetary-scale disturbances is best characterized by the flow of information from a zonal wavenumber-1 disturbance to a synoptic-scale circumglobal wave train whose amplitude peaks at the North Pacific and North Atlantic storm-track region. The feedback of synoptic-scale to planetary-scale disturbances manifests itself as a zonal wavenumber-2 structure driven by synoptic-eddy momentum fluxes. This wavenumber-2 structure locally enhances the East Asian trough and western Europe ridge of the wavenumber-1 planetary-scale disturbance that actively modulates the activity of synoptic-scale disturbances. The winter-mean amplitude of the actively interacting disturbances are characterized by pronounced fluctuations across interannual to decadal time scales.

Biology

Pacific Symposium on Biocomputing 7:498-509 (2002)

DISCOVERY OF CAUSAL RELATIONSHIPS IN A GENE-REGULATION PATHWAY FROM A MIXTURE OF EXPERIMENTAL AND OBSERVATIONAL DNA MICROARRAY DATA

C. YOO*, V. THORSSON♠, and G.F. COOPER*
*Center for Biomedical Informatics, University of Pittsburgh
8084 Forbes Tower, 200 Lothrop St., Pittsburgh PA 15213
♠The Institute for Systems Biology
4225 Roosevelt Way NE, Suite 200, Seattle Washington 98105

This paper reports the methods and results of a computer-based search for causal relationships in the gene-regulation pathway of galactose metabolism in the yeast *Saccharomyces cerevisiae*. The search uses recently published data from cDNA microarray experiments. A Bayesian method was applied to learn causal networks from a mixture of observational and experimental gene-expression data. The observational data were gene-expression levels obtained from unmanipulated "wild-type" cells. The experimental data were produced by deleting ("knocking out") genes and observing the expression levels of other genes. Causal relations predicted from the analysis on 36 galactose gene pairs are reported and compared with the known galactose pathway. Additional exploratory analyses are also reported.

Discovering Signaling Pathways – Karen Sachs

Get this book in print ▼

★★★★★
0 Reviews
Write review

Nanomedicine: A Systems Engineering Approach

edited by Mingjun Zhang, Ning Xi

Chapter Five

Learning Signaling Pathway Structures

Karen Sachs and Solomon Itani

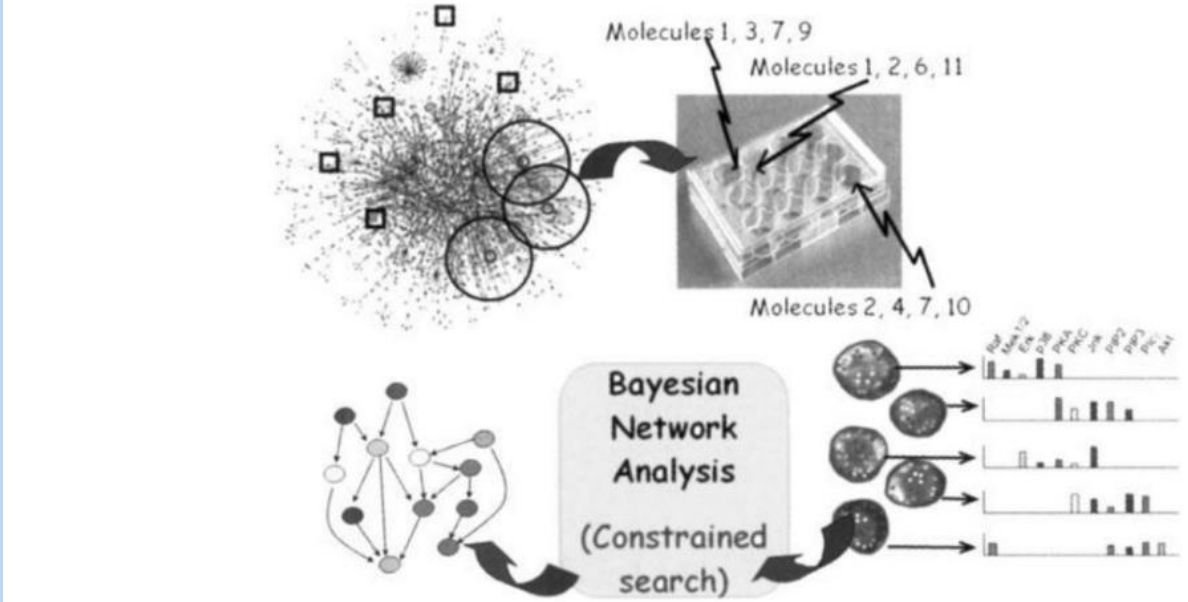


Figure 10. *Overview of Markov neighborhood algorithm.* A set of preliminary experiments are used to determine variable neighborhoods; these include variables which appear to be reasonable candidate parents based on a dependence metric (e.g. correlation). Sets of m molecules from each neighborhood are profiled under various stimulus conditions, and a constrained Bayesian network structure learning analysis is performed, in which candidate variable parents are selected from the (profiled portion of the) variable's neighborhood.

Discovering Signaling Pathways – Karen Sachs

Get this book in print ▼

★★★★★
0 Reviews
Write review

Nanomedicine: A Systems Engineering Approach

edited by Mingjun Zhang, Ning Xi

Chapter Five

Learning Signaling Pathway Structures

Karen Sachs and Solomon Itani

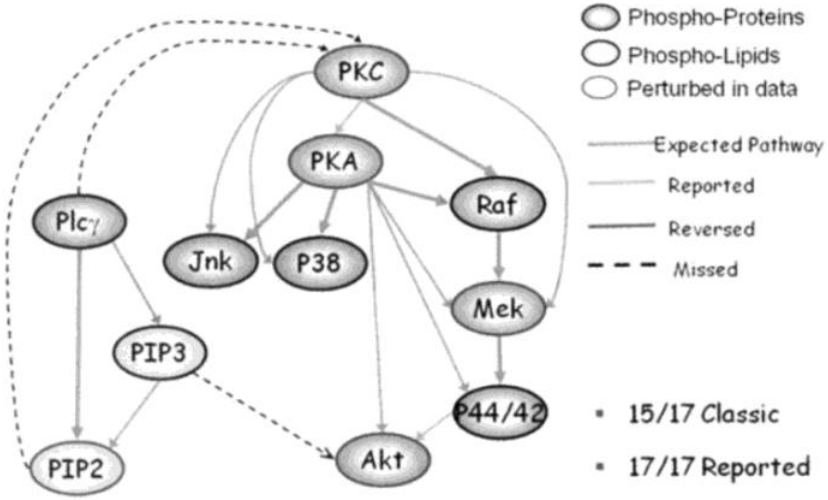


Figure 7. Model results contrasted to known influence connections. Literature reports were used to assess the retrieved network. "Classic/Expected" edges indicate edges that are well established in the literature, "Reported" edges are not well established and have never been reported in T-cells, "Reversed" indicates an edge which is directed incorrectly, and "Missed" indicates well established edges that were not found by the modeling approach.

Biology and Health

Psychological Medicine

cambridge.org/psm

Original Article

*Drs. Kummerfeld and Vinogradov contributed equally as senior authors to this manuscript.

Cite this article: Miley K, Meyer-Kalos P, Ma S, Bond DJ, Kummerfeld E, Vinogradov S (2023). Causal pathways to social and occupational functioning in the first episode of schizophrenia: uncovering unmet treatment needs. *Psychological Medicine* **53**, 2041–2049. https://doi.org/10.1017/S0033291721003780

Received: 10 April 2021
Revised: 30 August 2021
Accepted: 1 September 2021
First published online: 8 October 2021

Key words:
Early schizophrenia; motivation; social cognition; causal discovery; functional outcomes; machine learning

Author for correspondence:
Kathleen Miley, E-mail: mile0087@umn.edu

Causal pathways to social and occupational functioning in the first episode of schizophrenia: uncovering unmet treatment needs

Kathleen Miley[1,2], Piper Meyer-Kalos[1], Sisi Ma[3], David J. Bond[1], Erich Kummerfeld[3,*] and Sophia Vinogradov[1,*]

[1]Department of Psychiatry and Behavioral Sciences, University of Minnesota, Minneapolis, MN, USA; [2]School of Nursing, University of Minnesota, Minneapolis, MN, USA and [3]Institute for Health Informatics, University of Minnesota, Minneapolis, MN, USA

Abstract

Background. We aimed to identify unmet treatment needs for improving social and occupational functioning in early schizophrenia using a data-driven causal discovery analysis.

Methods. Demographic, clinical, and psychosocial measures were obtained for 276 participants from the Recovery After an Initial Schizophrenia Episode Early Treatment Program (RAISE-ETP) trial at baseline and 6-months, along with measures of social and occupational functioning from the Quality of Life Scale. The Greedy Fast Causal Inference algorithm was used to learn a partial ancestral graph modeling causal relationships across baseline variables and 6-month functioning. Effect sizes were estimated using a structural equation model. Results were validated in an independent dataset ($N = 187$).

Results. In the data-generated model, greater baseline socio-affective capacity was a cause of greater baseline motivation [Effect size (ES) = 0.77], and motivation was a cause of greater baseline social and occupational functioning (ES = 1.5 and 0.96, respectively), which in turn were causes of their own 6-month outcomes. Six-month motivation was also identified as a cause of occupational functioning (ES = 0.92). Cognitive impairment and duration of untreated psychosis were not direct causes of functioning at either timepoint. The graph for the validation dataset was less determinate, but otherwise supported the findings.

Conclusions. In our data-generated model, baseline socio-affective capacity and motivation are the most direct causes of occupational and social functioning 6 months after entering treatment in early schizophrenia. These findings indicate that socio-affective abilities and motivation are specific high-impact treatment needs that must be addressed in order to promote optimal social and occupational recovery.

Psychological Medicine

cambridge.org/psm

Original Article

Causal pathways to social and occupational functioning in the first episode of schizophrenia: uncovering unmet treatment needs

Kathleen Miley[1,2], Piper Meyer-Kalos[1], Sisi Ma[3], David J. Bond[1], Erich Kummerfeld[3,*] and Sophia Vinogradov[1,*]

*Drs. Kummerfeld and Vinogradov contributed equally as senior authors to this manuscript.

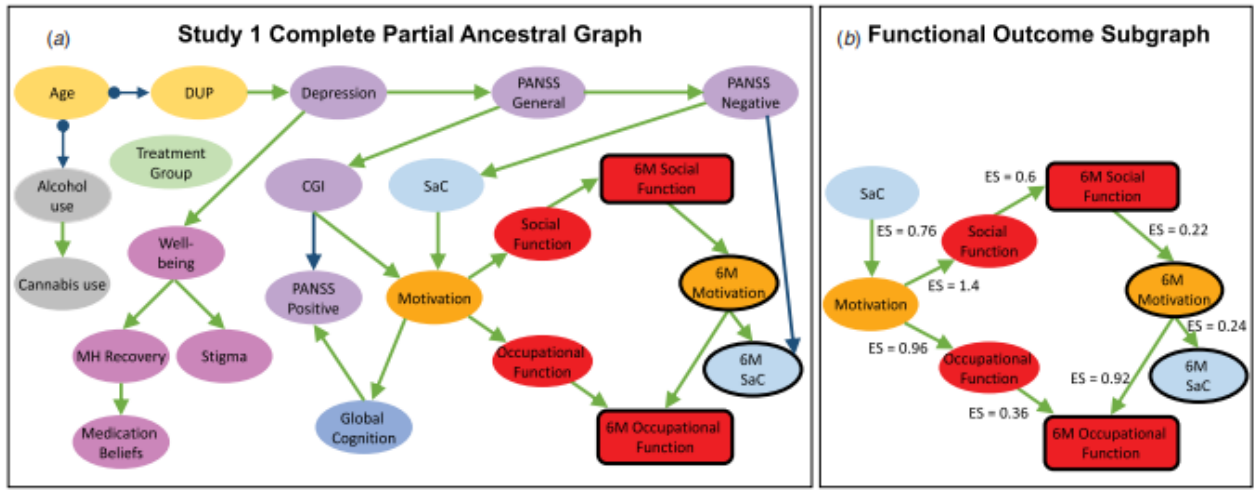


Fig. 2. Study 1 complete PAG and functional outcome subgraph. All variables represent baseline unless specified as 6M (6 month). Alcohol use = days of alcohol use

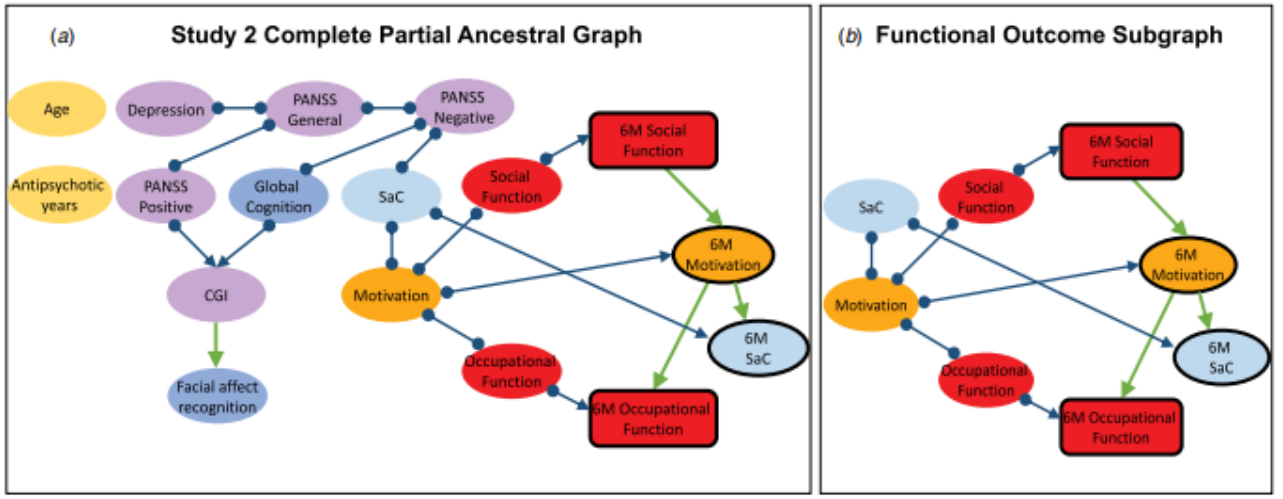
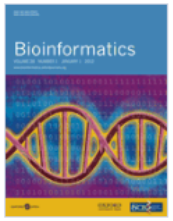


Fig. 4. Study 2: validation complete PAG and functional outcome subgraph. All variables represent baseline unless specified as 6M (6-month). Antipsychotic years = years of antipsychotic use; CGI = Clinical Global Impressions Severity score; Depression = Calgary Depression Scale for Schizophrenia; Facial Affect Recognition =



Volume 28, Issue 1
January 2012

Article Contents

< Previous Next >

JOURNAL ARTICLE

Inferring gene regulatory networks from gene expression data by path consistency algorithm based on conditional mutual information

Xiujun Zhang, Xing-Ming Zhao, Kun He, Le Lu, Yongwei Cao, Jingdong Liu, Jin-Kao Hao, Zhi-Ping Liu ✉, Luonan Chen ✉ Author Notes

Bioinformatics, Volume 28, Issue 1, January 2012, Pages 98–104,
https://doi.org/10.1093/bioinformatics/btr626
Published: 15 November 2011 **Article history** ▾

Abstract

Motivation: Reconstruction of gene regulatory networks (GRNs), which explicitly represent the causality of developmental or regulatory process, is of utmost interest and has become a challenging computational problem for understanding the complex regulatory mechanisms in cellular systems. However, all existing methods of inferring GRNs from gene expression profiles have their strengths and weaknesses. In particular, many properties of GRNs, such as topology sparseness and non-linear dependence, are generally in regulation mechanism but seldom are taken into account simultaneously in one computational method.

Results: In this work, we present a novel method for inferring GRNs from gene expression data considering the non-linear dependence and topological structure of GRNs by employing path consistency algorithm (PCA) based on conditional mutual information (CMI). In this algorithm, the conditional dependence between a pair of genes is represented by the CMI between them. With the general hypothesis of Gaussian distribution underlying gene expression data, CMI between a pair of genes is computed by a concise formula involving the covariance matrices of the related gene expression profiles. The method is validated on the benchmark GRNs from the DREAM challenge and the widely used SOS DNA repair network in *Escherichia coli*. The cross-validation results confirmed the effectiveness of our method (PCA-CMI), which outperforms significantly other previous methods. Besides its high accuracy, our method is able to distinguish direct (or causal) interactions from indirect associations.

SCIENTIFIC REPORTS
nature research

OPEN

Challenges and Opportunities with Causal Discovery Algorithms: Application to Alzheimer's Pathophysiology

Xinpeng Shen[1*], Sisi Ma[1], Prashanthi Vemuri[2], Gyorgy Simon[1*] & the Alzheimer's Disease Neuroimaging Initiative[†]

Causal Structure Discovery (CSD) is the problem of identifying causal relationships from large quantities of data through computational methods. With the limited ability of traditional association-based computational methods to discover causal relationships, CSD methodologies are gaining popularity. The goal of the study was to systematically examine whether (i) CSD methods can discover the known causal relationships from observational clinical data and (ii) to offer guidance to accurately discover known causal relationships. We used Alzheimer's disease (AD), a complex progressive disease, as a model because the well-established evidence provides a "gold-standard" causal graph for evaluation. We evaluated two CSD methods, Fast Causal Inference (FCI) and Fast Greedy Equivalence Search (FGES) in their ability to discover this structure from data collected by the Alzheimer's Disease Neuroimaging Initiative (ADNI). We used structural equation models (which is not designed for CSD) as control. We applied these methods under three scenarios defined by increasing amounts of background knowledge provided to the methods. The methods were evaluated by comparing the resulting causal relationships with the "gold standard" graph that was constructed from literature. Dedicated CSD methods managed to discover graphs that nearly coincided with the gold standard. For best results, CSD algorithms should be used with longitudinal data providing as much prior knowledge as possible.

Summary / Conclusions

A lot of good work on important scientific questions -

- But distributed across disciplines, so largely unknown
- Almost all of it by algorithm builders/developers
- Too little of it includes experimental follow-up

By far the most common use is mechanism discovery

We need a sub-discipline devoted to applications