Manipulability theories: Overview

**Basic idea:** $A$ causes $B$ if bringing about $A$ is an effective way to bring about $B$.

**Elaboration** (Woodward): there is a stable or invariant relationship between $A$ and $B$ that can be used to show how $B$ would change by interventions that change $A$.

**Alleged virtues:**
- improves on counterfactual account. The relevant counterfactuals have an antecedent that is brought about by an *intervention*, at least for basic cases of causation (cases on a ‘human’ scale).
  - less problematic truth conditions than general counterfactuals.
  - retains most advantages of the counterfactual account (solution to causal asymmetry)
  - good handling of spurious causation (e.g., barometer/storm)
  - good handling of pre-emption cases (desert traveler; trumping)
  - gives a definite answer in over-determination cases (each bullet counts as a cause) [virtue or vice?]
- moves away from observation-based metaphysics/epistemology [virtue or vice?]

**Disadvantages/Objections:**
- confuses metaphysics (what causation *is*) with epistemology (how we infer causal connections)
- apparent circularity
  “Bringing about A” or intervening is itself a causal notion.
- antropocentrism
  If there were no humans (agents), would there be no causation?
  If humans had different powers of intervention, would causation be different?
- limited scope
  There can be no causation where (human) intervention is impossible (e.g., supernovae).
  “$A$ causes $B$” is false whenever $A$ cannot be brought about through (human) manipulation.
I. von Wright: On the Logic and Epistemology of the Causal Relation

A model to see how far logic can go in capturing the nature of causal relationships: though the non-logical sections (7-9) have been much more influential.

Limited to deterministic causation.

2. Logical analyses appear unable to account for causal asymmetry

Causal claims are relations between (the obtaining of) states of affairs $p$ and $q$

- This connection can be construed as sufficiency or as necessity
  [N.B. an early effort that does not discuss partial causes]
- The analysis of sufficiency or necessity may be extensional or intensional

  a) Extensional: universal quantification over times.

     Whenever $p$ obtains, $q$ obtains. (Sufficient condition)
     Whenever $q$ obtains, $p$ must have obtained. (Necessary condition)

     Either is a version of the Humean ‘regularity’ analysis (as are Mackie’s INUS conditions); necessity and sufficiency are analyzed purely in terms of what goes on in the actual world.

  b) Intensional: modal analysis (= quantification over all possible worlds or states of affairs, including all times in each possible world)

     In all possible worlds: if $p$ obtains, $q$ obtains. (Sufficient condition)
     In all possible worlds: if $q$ obtains, $p$ must have obtained. (Necessary condition)

     [Lewis’ counterfactual analysis counts as intensional as well, in a broad sense. We have to ‘consult’ worlds other than the actual one to determine the truth-values of causal claims.]

- The extensional analysis is too weak.

  Argument: Accidental uniformities do not support counterfactuals. But causal claims must support counterfactuals.

  [N.B. He does not address the objection that his version of the intensional analysis is clearly too strong. Lewis’ Nixon example shows this.]

- The Problem: Neither the extensional nor intensional analysis accounts for the asymmetry of cause and effect. The logical relations can be contraposed, but not the causal ones. We don’t say that non-flooding causes the absence of rain.

- Rejects accounts that build in temporal priority (because of simultaneous causation)
3.4. **Logical preliminaries**

- Start with propositional logic
- Add a primitive modal operator $M$ (possibility) and derived operator $N$ (necessity)
  
  $$Np \equiv \neg M\neg p$$

- Add primitive tense-logic operators $T$ (‘and next’) and $\lor$ (‘at some future time’) and derived operator $\land$ (‘at all future times’)
  
  $$\land p \equiv \neg \lor \neg p$$

- **Background framework** of branching time: a ‘tree’ of discrete moments
  - Each node is a complete state of the world at one moment
  - One branch into the future for each possible next state
  - One and only one of these future states *will* obtain
  - No backward branching
  - A *complete* chain of nodes (from beginning to end) is called a *history*

- **Simplifying assumptions**

  *Logical Atomism.* A complete state can be specified by specifying the value (T or F) of $n$ independent (atomic) binary variables. Hence there are $2^n$ possible states. (cf. Keynes’ *Limitation of Independent Variety*)

  *Finite time.* The total number of nodes in any history is $m$.

  Hence, if what happens at the next node is totally independent of the preceding node (i.e., no causality), there are $2^{mn}$ possible histories.

- At a given node, if only $k$ alternatives emerge, the degree of freedom is

  $$(\text{# of causally possible alternatives} - 1) = k - 1$$

  $$(\text{# of logically possible alternatives} - 1) = \frac{2^n}{2} - 1$$

  **Key idea:** causation is a constraint on what’s possible
5. Semantics

All truth-values are assigned relative to a possible world or state of affairs.

Propositional calculus
(1) Atomic propositions are assigned T or F at each world.
(2) Truth-functional compounds of atomics get truth-values using truth-tables.

Modal operator
(3) $Mp$ has two possible analyses:
   (i) Immediately possible. $Mp$ at $w$ if $p$ holds at one of the worlds possible immediately after $w$;
   (ii) Eventually possible. $Mp$ at $w$ if $p$ holds at one of the worlds eventually possible in the future after $w$.

   In neither case is “actual implies possible” ($p \to Mp$) valid.

Tense-logic
(4) $pTq$ is true at $w$ iff $p$ holds at $w$ and $q$ holds in the world that will come true next.
(5) $\lor p$ is true at $w$ iff $p$ holds at some world that will come true in the future

Note that if $t$ is any tautology, $tTq$ at $w$ really just says: $q$ holds in the world that will come true next.

The difference between the two accounts of $Mp$:

   $tTp \to Mp$ on either first or second account
   $\lor p \to Mp$ only on second account

von Wright takes the second account to be more important.

   The S4-like principle $MMp \to Mp$ is valid only for the second semantics.
   The principle $Np \to \land p$ is valid only for the second semantics.
6. Analysis of causal relation

Case 1: p is a sufficient condition for q

All of these analyses use the second, S4-like semantics for M. [It’s not even clear why the first was needed.]

a) Whenever p happens, q immediately follows.
   \[ N(p \rightarrow Tq) \]

b) Whenever p happens, q happens simultaneously.
   \[ N(p \rightarrow q) \]

c) Whenever p happens, q follows at some later time.
   \[ N(p \rightarrow \nu q) \]

[**Puzzle:** Even on the second reading of M, these claims will not apply to the present moment or ‘world’, but only to future worlds. Here it’s odd that we don’t have \( Np \rightarrow p \).]

Case 2: p is a necessary condition for q

a) p was there immediately prior to q.
   \[ N(Tq \rightarrow p) \]

   [Implies not-p is a sufficient condition for not-q.]
   [von Wright’s remark at end of §6 is misleading. He departs from the definition he has just given of necessary condition!]

b) p is there at the same time as q
   \[ N(q \rightarrow q) \]

c) p was there at some previous time.
   ??

**N.B.** This could be extended to INUS conditions. Apply this to the total necessary and sufficient condition of which the INUS condition is a part. (But since it is only a model, and obviously a vast simplification, there is little point.)
7. Epistemology of the causal relation

**Problem:** How can we know that \( p \) is a sufficient cause of \( q \)?

Restrict to \( N(p \rightarrow Tq) \): necessarily, when \( p \) happens, \( q \) immediately follows. This entails:

i) \( \land (p \rightarrow Tq) \): whenever \( p \) happens, \( q \) immediately follows.

ii) When \( p \) does not obtain, \( q \) would have followed had \( p \) obtained.

ii) distinguishes a causal connection from a universal regularity. But not even i) can be known with certainty (Problem of Induction). What justifies i) and ii)?

**Ideal:** To test the counterfactual, substitute for a world in which \( p \) is not the case another world where \( p \) is the case, to show that \( q \) would have followed. We can’t do this in the present or the past, but we can in the future through *action*.

**Re-interpretation of the branching tree:**

Top branch out of each node reflects what the ‘natural’ development would be if nature is *left alone*.

Other branches reflect what happens if *agents* act.

Pre-supposition of action: things often continue as they are unless we act.

**Two crucial ‘experiments’:**

1) Make a \( p \)-world where it would have been not-\( p \).

   If \( q \) follows, this “verifies” the counterfactual \( p \rightarrow q \).

   It also rules out the possibility that regular succession from \( p \) to \( q \) is due to a common cause \( c \): since this would have been not-\( p \) but for our action, no such common cause can be present.

2) Abstain from interfering with a not-\( p \) world.

   If \( q \) does not follow, this “verifies” the counterfactual \( \neg p \rightarrow \neg q \).

   It also rules out the possibility that the regular succession from \( p \) to \( q \) is due to some remote cause \( c \) of \( q \) (*spurious causation*): otherwise, \( q \) would have happened despite \( p \) not happening.

**Upshot:** What confers on observed regularities the character of a causal connection is the possibility of subjecting the causal factor to experimental tests by interfering with the natural course of events.

**Comment:** for von Wright, the dependence upon action is *epistemological* and not ontological, but the logical characteristics of causation and action are closely connected.

- “Verification” is in scare quotes: full verification is impossible
- causation need not always result from action, but test-procedures result from action alone (scientific experiments)
8. Asymmetry

Asymmetry between cause and effect arises if we bring about changes in $q$ by manipulating $p$, but not vice versa.

**Advantage:** manipulability accounts can explain causal priority in cases of simultaneous causation.

Note that both could be causes and effects (card house)

**Functional laws**

Ex: Ideal Gas law, $PV = kT$

i) Express causal relationships, but only in particular cases *when* manipulation takes place can one distinguish causes from effects.

There is no cause-effect relationship at the type level.

ii) Non-manipulable factors can only ever be effects.
9. Pre-suppositions for manipulative causation

1) Recurrence of identical states (or partial states); knowledge of how the world will develop without interference.
2) Logical atomism: independence of cause and effect.
3) Familiarity with regularities.

Objection: Causal relations between phenomena not subject to interference can’t be accounted for this way.

Reply:

i) At times, these relations involve non-causal laws.
ii) At times, they are derived from other laws that have been tested, and hence may be considered causal laws.
iii) At times, the laws are really part of our conceptual framework (e.g., conservation laws).

[Note: More thorough responses to this objection in Woodward and in Menzies/Price]

10. Applicability to human sciences

Problem: Explanations involving reasons. If reasons are causes, then we lose the independence of cause and effect that we find in the natural sciences.

Thesis: the existence of reasons does not cut down on alternate possibilities, since it never necessitates that the agent will not change his/her mind. The connection between reasons and action is logical, not causal.

Limitation on treating human behaviour causally: the notion of cause depends on a sui generis (non-causal) conception of agency and action.
II. Menzies and Price: Causation as a Secondary Quality

A sustained defense of the manipulability or agency theory of causation:

1) Major advantages over rival accounts:
   1. nice treatment of probabilistic causation (unified account of deterministic and probabilistic causation)
   2. avoids problem of spurious causation
   3. moves away from ‘spectator’, observation-based empiricism

2) Stock objections are mistaken:
   1. Confusion of epistemology of causation with its metaphysics: manipulation is evidence for causal claims, but not part of their meaning.
   2. Circularity: intervention or bringing about is itself causal.
   3. Limited scope: cannot make sense of causal relations between events outside our powers of intervention.
   4. Anthropocentric: the account implies that causal relations would not exist, or would be different, if humans did not exist or were different.

Strategy: parallel objections to human-dependent accounts of colour all fail. In a suitably extended sense, causation is a secondary quality like colour.
2. **Agency theory**

a) **Manipulability theories:** $A$ is a cause of $B$ if bringing about $A$ would be an effective means of bringing about $B$.

b) **Probabilistic causation:** $A$ causes $B$ if $A$ raises the probability of the occurrence of $B$. 

M&P combine a) and b) in their agency theory of causation.

- **Agent probabilities.** $P(B \mid A)$, or $P_A(B)$, where $B$ is an event (state of affairs? proposition?) and $A$ is something that can be realized by the agent’s action.
  
  [Suggestion: think of $A$ as an agentive sentence, and $B$ as a sentence describing the relevant state of affairs.]

- Characterized *functionally* as the probabilities to be used in *decision theory*.

  Ex: Unilateral disarmament
  Ex: Medical Newcomb case
  Ex: Newcomb’s problem

Evidential probabilities vs. causal probabilities (propensities)
These probabilities are *subjective*, but that does not mean they are not based on evidence.

**Definition:** $A$ causes $B$  
iff $A$ is an effective means of achieving $B$  
iff $P_A(B) > P_{\neg A}(B)$.

**Gloss:** Given that $B$ is desired more than $\neg B$ in all circumstances, decision theory prescribes that you should do $A$. This is the meaning of $A$ is an effective means of bringing about $B$.

**Comments**

1. Neither $A$ nor $\neg A$ may be particularly effective (both probabilities may be low), but so long as $A$ is *more effective* than $\neg A$ in bringing about $B$, it is a cause of $B$.

2. Unstated: we suppose that $A$ and $B$ have actually occurred.
Virtues

1. Nice theory of probabilistic causation, with deterministic limiting case (probability 1).
2. Avoids spurious causation.
   - With *evidential* conditional probabilities, $A$ raises the probability of $B$ even if $A$ is caused by $B$, or if both $A$ and $B$ are effects of a common cause $C$.
   - This won’t happen with agent probabilities, which are *not* evidential.

Example: $A =$ falling barometer, $B =$ storm ($C =$ falling air pressure)

\[
P(B \mid A) > P(B \mid \sim A) \text{ with evidential probabilities}
\]

\[
P_A(B) = P(B) = P_{\sim A}(B) \text{ with agent probabilities, so there is no causal connection.}
\]

To compute the first probability, you create a **brand new causal history** that does not assume $C$ (or $\sim C$), but involves directly manipulating (or not) the barometer.

3. Makes intervention a key (irreducible) feature of the world.

   Observation-oriented empiricism has no way to handle spurious causation
   In fact, the way we do handle it is with experiments/interventions

   [Qu: making intervention a part of *epistemology* is a good idea; but metaphysics?]
3. Objection 1: confuses metaphysics and epistemology

Objection: Experiments and interventions are the main tool for discovering and testing causal claims, but there is no reason to think they are part of the meaning of causation.

cf. Verificationism: the logical empiricist doctrine that the meaning of a statement is its method of verification.

- failed for many reasons (see Hempel, “Empiricist Criteria of Meaning…”)

Response: The very notion of causation is rooted in the idea of manipulation.

Colour analogy: To be red is to look red to a normal observer under standard conditions.

- Takes ‘looks red’ to be primary
- Defines ‘is red’ only in relation to a human observer (secondary quality)
- Epistemological implications, but the analysis is pure metaphysics

Causal relation: For A to be a cause of B is for \( P_A(B) > P_{-A}(B) \), i.e., for A to be an effective way to bring about B for a (normal?) agent (under standard conditions?).

- Takes ‘brings about’ to be primary [Problem: this is not so obviously in a different ontological category from ‘causing’, as ‘looks red’ is in a different category from ‘is red’]
- Defines ‘cause’ in relation to a human agent (see anthropomorphism objection)
- Epistemological implications (e.g., possibility of error), but analysis is pure metaphysics
4. Circularity

**Recall:** A causes B if bringing about A would be an effective means by which an agent could bring about B.

**Objection:** Bringing about is a causal notion, so the agency theory can’t be a reductive analysis of causation.

**Response 1 (informative circularity):** The account is not reductive, but nevertheless non-trivial. (It provides informative linkage between causation in general and the special case of agent causation.)

**Response 2 (agency is sui generis):** Bringing about is a basic kind of relation that is prior to any notion of causality.

*Qu:* Prior in what sense? Epistemologically, maybe. But why metaphysically?

**Colour analogy:** “Is red” is analysed in terms of “looks red”, but here circularity is no problem.

*Qu:* Isn’t there a disanalogy here? “Looks red” is not a species of being red, but bringing about is a case of causation.

I don’t think the analogy is much help.

A closer look at Response 2:

- we get the notion of bringing about from ostension or acquaintance with our own successful bodily movements that achieve our ends

- the general notion arises from these basic cases as in the definition

**Some random reflections:**

The basic cases have to be construed as non-causal relations, rather than as paradigms of causation.

They don’t fit the definition (ex: thumping my fist). But perhaps that is a problem with the definition.

Are the basic cases (‘bringing about A’) non-relational – a single event, rather than a relation of events?

Why do M&P suddenly shift gears to epistemology, right after worrying about confusion between metaphysics and epistemology?
5. Unmanipulable causes

Objection: The definition, strictly speaking, limits the scope of the causal relation to pairs $A-B$ where $A$ is something that can be (humanly) manipulated. So causation would be inapplicable to earthquakes, cosmic events, etc., which is absurd.

*Colour analogy:* Colour claims, strictly speaking, are inapplicable to cases where it is impossible for a normal observer to observe under standard conditions.
(Ex: inside the sun)

Response 1: The relevant counterfactuals are still true (even though the antecedents are physically impossible to realize).
- move to idealized observer (colour case) or agent (causation case)
- rejected: a) *finkish* dispositions (vanish when put to the test)
  b) *swamped* dispositions

Response 2: *Quasi-dispositions.*

*Colour analogy:* The disposition to ‘look red’ is produced by *intrinsic properties* (even in cases of finkish dispositions, or swamped dispositions).

Objects with the identical intrinsic properties (but not the other masking properties) *would* look red.
So by analogy, we can say these problematic objects have the quasi-disposition to look red, and hence (on an extended definition) *are* red.

$x$ is red iff $x$ would look red to a normal observer… or $x$ possesses intrinsic properties identical or similar to those of an object $y$ that would look red…

*Causation case:*

$A$ causes $B$ iff bringing about $A$ would be an effective way for an agent to bring about $b$, or the “situation” possesses intrinsic features similar to those of another “situation” involving an analogous pair $A', B'$ of which the counterfactual does hold.

Comments:

1) The analogy is much vaguer in the causation case than in the colour case.
   An artificial simulation of an earthquake is good enough.
   But why not an artificial simulation of a weather system, in which you can ‘bring about’ an artificial storm by fiddling with the barometer? Aren’t the items similar enough?
   Point: new causal relations may emerge in the model, and old ones be omitted. Need case-by-case analysis, not a blanket argument.

2) Lewis rejects this approach in his “Causal Influences” paper.
   Unable to handle *trumping* cases.
6. Anthropocentrism

*Naïve objection:* The agent theory entails that there are no causal relations if there are no agents (or no manipulations).

*Response:* The account is *counterfactual* and does not require actual manipulations.

*Colour analogy:* the tomato is still red even if nobody looks at it.

*Sophisticated objection:* In worlds where agents have different powers, causal relations are different. This *relativizes* causal claims to agents’ powers.

*Ex:* if our powers of manipulation were different, there would be fewer causal relations.

*Colour analogy:* in worlds where observers are colour-blind, there is no distinction between red and green (say).

*Response 1 (rigidity):* the disposition to look red is defined rigidly, relative to normal observers in the actual world. Similarly, the relevant powers of manipulation for defining causal relations in any world $w$ are the powers of agents in the actual world.

This just appears dogmatic. The agents in $w$ would have the same conception of agency as we do, but would take the scope of the causal relation to be much more limited. How do we know we are not in their position with respect to some other world where agents have even greater powers?

*Response 2 (insensitivity):* Accept the relativization, but note that it is innocuous.

i) In the case of colour, relativization is unproblematic. (Lemons *would be* sweet if our taste buds were different.) Why can’t this be the case for causation too?

ii) The *extended* definition of causation makes the relation relatively independent of our actual powers. (Cf. computability)

   Exception: “intelligent trees”.

   *[Qu: M&P say they would have ‘no notion’ of causation; but they should accept that there would be no causation.]*

*Generalized* notion of secondary quality: should not be restricted to monadic properties, and should not be restricted to observers.
III. Woodward: “Explanation, Invariance and Intervention”

- Equates causation and explanation at type level: \(A\) explains \(B\) is synonymous with \(A\) causes \(B\) (at type level)
- Explanatory relations are *invariant relations* that support *interventions*

**Invariant relations**: a generalization of laws of nature.
**Interventions**: experimental manipulations; a basic notion.

*Background:*
1) D-N model of explanation
2) Two weaknesses:
   - essential inclusion of a *law of nature* (a problem for sciences with few laws)
   - problems of *asymmetry* and *irrelevance* (traced to neglect of causation)

Invariant relations address the first weakness; interventions address the second weakness.

2. *Explanation*

*Example 1:* Derivation of the electric field of an infinite wire of uniform charge distribution \(\lambda\) at a distance \(r\) from the wire.

Need to distinguish between the *narrow* explanation (the deduction using these particular values) and the *broad* explanatory pattern (the style of argument or pattern of derivation).

- **Narrow explanation** is textbook D-N: a valid derivation involving Coulomb’s Law
- **Broad explanation** answers counterfactual questions about how the *explanandum* would have changed had the initial and boundary conditions been altered: changes in density \(\lambda\), or distance \(r\), or in the geometry of the conductor

- the law itself is *invariant* under these alterations
- the relevant alterations are *interventions*
- significance of the broader explanatory pattern:
  i) we see what the electric field depends upon
  ii) we can solve the problems of irrelevance (hexed salt, birth control pills) and spurious causation (barometer/storm)
  iii) we can solve the problem of asymmetries (flagpole/shadow)
3. Counterfactuals and Interventions

The broad explanatory relation is linked to the ability of an explanatory account to provide counterfactual information.

1. How are counterfactuals interpreted?
2. Which ones are important?

1: The relevant counterfactuals for explanations (and hence, for causal claims) are those whose antecedents are made true by external interventions.

*Intervention* = experiment or manipulation that can be carried out by an idealized human

Roughly: the interpretation is that \( A \rightarrow B \) is true iff the intervention that brings about \( A \) has ‘the right causal history’: is entirely responsible for the change to \( B \) and the intervention is independent of all other possible causes of \( B \). (Cf. M&P: start an independent history)

It is the interpretation under which joint effects of a common cause are not counterfactually dependent upon each other (barometer/storm)

Two objections:

1) Human-dependent (M&P’s *scope* and *anthropocentrism* worry).

*Response*: the account (M) of interventions on p. S30 is *not* specifically about human interventions. A natural process may count as an intervention.

2) Circularity.

*Response*: A non-reductive account, but informative: analyzes causal claims using interventions. The interventions are processes with the ‘right causal history’.
4. Invariance

A generalization is *invariant* under a class of interventions if it continues to hold under these interventions.

In Example 1, the invariant generalization is Coulomb’s Law.

Explanations are derivations that appeal to generalizations that are invariant under *some class* of interventions that includes, at minimum, changes to the initial or boundary conditions.

- That explanatory relations must be invariant is common in many sciences.
- In special sciences, one finds many invariant relations that are not laws.

5. An example

[‘Econometrics’ is a misprint]

Consider an equation such as:

\[ Y = a_1X_1 + a_2X_2 + \ldots + a_nX_n + U \]

- \( Y \): dependent variable (e.g., plant height)
- \( X_i \)’s: independent variables (e.g., amount of water, fertilizer)
- \( U \): error term
- \( a_i \)’s: coefficients to be determined

(1) If the equation represents a causal or explanatory relationship, this *means* that over a certain range of values for the \( X_i \)’s, if one of the dependent variables \( X_i \) is changed by \( \Delta X_i \), then \( Y \) should change by \( a_i \Delta X_i \). The equation is invariant over some range of interventions that modify the \( X_i \)’s, and supports counterfactual claims in this range.

(2) If the equation is *not* invariant over any range of interventions, then it is a mere statistical relationship among these particular values and does not describe an explanatory relation.

**Invariant relations vs. laws of nature:**

Laws of nature are (ideally) exceptionless and universal: hence, invariant over all interventions. (But in fact, it’s a matter of degree.)

A relationship such as the equation can be invariant over some range of interventions, but not over all (e.g., temp. 1000 °C), and hence *not* lawlike.

The main thesis of the paper is that, within its domain of invariance, a generalization can be used in explanations.
6. Invariance and Explanatory Depth

Wider ranges of invariance ⇒ deeper explanations.

Graphical representation of causal relations:
- each node is an event (or event type)
- arrows represent direct causal connections
- a node can only causally affect other nodes that can be reached by forward arrows

Equations correspond to graphs as illustrated (p. S35). Nodes with no ‘in’ arrows are independent or exogenous variables.

• A system of equations has different possible graphical representations
• Many such representations are generally consistent with observed data, but they differ on counterfactual implications

If equations express causal (explanatory) relations, they exhibit two types of invariance:

    i) \textit{functional form invariance}: functional form does not change under changes to variable values
    ii) \textit{coefficient invariance}: the coefficients can each be changed separately by interventions that have no effect on other coefficients. (\textit{Modularity} of causal relationships).

Although many equations are consistent with observed data, only one set of equations will satisfy functional form and coefficient invariance. (Mechanistic picture of explanation: machines have this characteristic modularity.)

Coefficient invariance is also entailed by M, in some contexts.

Comparative judgements of the range or degree of invariance are based on range of interventions over which equations are invariant.
7. Robustness and invariance

Robustness of probabilistic causes: $A$ is a robust probabilistic cause of $B$ if

\[ P(B / A) = P(B / A \cdot D_i) \text{ and } P(B / \neg A) = P(B / \neg A \cdot D_i) \]

for different ways $D_i$ of bringing about $A$.

Relative to $A$, $B$ is independent of each $D_i$. Or: $A$ screens off $B$ from $D_i$.

Problem: This has no counterfactual import; it is defined in terms of actual statistics. Different systems of equations can be compatible with observed data and obey the very same screening off relationships, but yet have different counterfactual implications.

There are other problems with robustness that can be met with invariance.