Where is the Theory in our ‘Theories’ of Causality?

Nancy Cartwright
LSE and UCSD

Forthcoming in Hunting Causes and Using Them, CUP and the Journal of Philosophy
'Theories' of Causality

- The probabilistic theory of causality (Patrick Suppes) and its descendents
  - Bayes-nets theories (Wolfgang Spohn, Judea Pearl, Clark Glymour)
  - Granger causality (economist Clive Granger)
- Modularity accounts (Pearl, James Woodward, economist Stephen LeRoy)
- Manipulation accounts (Peter Menzies, Huw Price)
- Invariance accounts (Woodward, economist/philosopher Kevin Hoover, economist David Hendry)
- Natural experiments (Herbert Simon, economist James Hamilton)
- Causal process theories (Wesley Salmon, Philip Dowe)
- The efficacy account (Kevin Hoover)
- Counterfactual accounts (David Lewis, Hendry, social scientists Paul Holland and Donald Rubin)
1. Causation: One Word, Many Things
2. Causality: Metaphysics and Methods
3. Representation: Handsome is as Handsome Does
4. Causal Laws and Effective Strategies
5. Hunting Causes v Using Them
6. Causality: Metaphysics, Methods and Use
7. Where is the Theory in our ‘Theories’ of Causality?
1. Causation: One Word, Many Things
A *Bayes-net account* of what causality is, of the kind Wolfgang Spohn offers, is provably correct for a system of causal laws over a causally sufficient set of quantities for which the three Bayes-nets postulates (causal Markov, minimality and faithfulness) hold plus the assumption that if any two factors are probabilistically dependent then either one of them causes the other or they have a common cause.

Similarly I prove a representation theorem that shows that for systems of laws that satisfy the axioms of a ‘linear deterministic system’, *Woodward’s invariance account* of causation always yields correct results.
2. Causality: 
Metaphysics and Methods
Probabilistic ‘Theory’ of Causation

THEORY:
C causes E in $K_i$ iff
$P(E/C&K_i) > P(E/\neg C&K_i)$
($K_i$ is a state description over a complete’ set of confounding factors)

METHOD:
‘C causes E’ is licensed if C and E are probabilistically dependent once we have stratified on a ‘complete’ set of confounding factors.
3. Representation: Handsome is as Handsome Does

Rather than looking for one special relation in the world that legitimates representing them as causal, look instead for some unified features of the representations themselves.

TRY: A scientific representation should allow us to make inferences about the system represented.
4. *Causal Laws and Effective Strategies*

Consider one kind of inference that has always been stressed as central to the notion of causation: representing a set of relations as causal should allow us to make some kinds of inferences that allow us to use causes as strategies for producing effects.
Kevin Hoover:

$x$ causes $y$ iff anything we can do to fix $x$ affects $y$ but not the reverse.

If that’s what causality is, then the inference from ‘$x$ causes $y$’ to ‘We can influence $y$ by fixing $x$’ is built right into the metaphysics.
David Hendry:

- Causes are superexogeneous.

- Superexogeneity = weak exogeneity + ‘invariance’.
**Weak Exogeneity:**
Given \( P(y \& x, \beta U \gamma) = P(y|x,\beta)P(x,\gamma) \)

\( x \) is exogeneous to \( y \) if the parameters \( \gamma \) of the marginal distribution of \( x \) have no cross-restraints with the parameters \( \beta \) of the conditional distribution of \( y \) on \( x \).

**Superexogeneity:**
In addition the conditional distribution is invariant to interventions that change (the parameters of) the marginal distribution.

Again an inference about use follows immediately from the very notion of causation.
5. Hunting Causes vs Using Them
Consider *Woodward*:

Given $y = f(x,\ldots)$, $x$ causes $y$ iff

1. modularity: $x$ can be varied independently of all ‘other’ sources of variation in $y$
2. invariance: $y = f(x,\ldots)$ continues to hold as $x$ is varied holding fixed other sources of variation in $y$

Woodward’s theory of causality licenses counterfactuals about what would happen if the cause were to vary in a very special way, a way that is not what we would normally be envisaging for either policy or technology. The variation is just the kind we need if we want to test a causal claim – to hunt for causes – but not the kind we expect to implement when we try to use them.
Consider our list of causal theories. They are in essence either 'language-entry rules' – they tell us when it is correct to label relations 'causal' (and generally only one method for so doing) – or they are 'language exit rules' that tell us what inferences we can draw from that label (and again, usually only one particular kind of inference). This raises the twin concerns

- What is the use of causes characterized from a hunting methodology?
- How do we hunt causes when they are characterized in terms of their use?
6. Causality: Metaphysics, Methods and Use

Our philosophical treatment of causation must make clear why the methods we use for testing causal claims provide good warrant for the uses to which we put those claims.
7. *Where is the Theory in our ‘Theories’ of Causality?*

For manipulation it seems that any kind of formula that allows us to predict what we want to know and that holds true across the variations we envisage making will do.
The work in deriving the relevant results is done *not* by the axioms of the system but by the very restricted notion of ‘manipulation’ that is used. The manipulations involved in these theorems are like Woodward’s (or those miracle manipulations of David Lewis) – the kinds of manipulations we should like to make in a Galilean experiment, not the kind we envisage for policy, where many things may be expected to change at once, including some of the causal principles by which the system operates.
All cases where counterfactual inferences or manipulation theorems are proven from some system of causal principles work like this:

First, a set of relations is postulated that satisfies some set of constraints taken to be sufficient to characterize the relations as ‘causal relations’.

Second, manipulations are taken to create a new, closely related system. The new systems are defined by essentially two clauses.

- Describing what is to change, e.g. the value of certain variables, maybe representing quantities all of which can vary independently of each other (i.e. they are ‘variation-free’), or the relation by which the ‘manipulated’ quantity is produced.

- Stating a set of things that do not change, like values of other particular variables and all the other specified relations except that by which the ‘manipulated’ quantity is produced.

Third, the counterfactual consequence is calculated in the new system.
NOTICE: the constraints that justify labelling the system of relations as 'causal' have no role to play. What matters is that there is a system of relations, say equations, that allow us to solve for a given result. Then we are given a recipe for making very specific changes and the recipe produces a new system that also allows us to solve for the result in question. The solvability of the two systems plus the rule for getting from one to the other is what does the job, not any of the constraints that make it a 'causal' system in the first place (beyond those that ensure solubility).
BUT>>>>

Maybe our whole enterprise is misplaced. Perhaps we should be seeking not theories of causality but rather causal theories of the world.