Extrapolation, Analogy, and Comparative Process Tracing

Francesco Guala[†‡]

Comparative process tracing is the best analysis of extrapolation inferences in the philosophical and scientific literature so far. In this essay I examine some similarities and differences between comparative process tracing and former attempts to capture the logic of extrapolation, such as the analogical approach. I show that these accounts are not different in spirit, although comparative process tracing supersedes previous proposals in terms of analytical detail. I also examine some qualms about the possibility of drawing extrapolation inferences in the social sciences and conclude by suggesting that there may be cases of extrapolation without process tracing.

1. Introduction. The problem of extrapolation (or external validity, as it is sometimes called) is a minor scandal in the philosophy of science. Ever since Galileo and Newton, it has been agreed generally that experiments constitute the ideal, most reliable source of scientific knowledge. And yet, until recently philosophers lacked a general, convincing account of how such knowledge—gained in special, highly controlled circumstances—could be applied in circumstances where the special conditions do not hold. Science, as usual, was ahead of philosophy: specific approaches to extrapolation could be found in various disciplines, such as the distinction between phase 1, phase 2, and phase 3 trials in medical research.[1] Still,

†To contact the author, please write to: Department of Economics, University of Milan, Milan, Italy; e-mail: francesco.guala@unimi.it.

‡Earlier versions of this essay were presented at the 2008 International Philosophy of Science conference in Dubrovnik and the 2008 Philosophy of Science Association meeting in Pittsburgh. I thank participants to both conferences and especially Julian Reiss and Daniel Steel for detailed comments that helped me during revision. The usual caveats apply.

1. To receive FDA approval, drugs must be tested on animals first, then on human beings in "ideal" experimental settings, and finally must pass so-called efficacy trials with patients in real-world conditions.

Philosophy of Science, 77 (December 2010) pp. 1070–1082. 0031-8248/2010/7705-0007$10.00
Copyright 2010 by the Philosophy of Science Association. All rights reserved.

https://doi.org/10.1086/656541 Published online by Cambridge University Press

because such methods are tailored to the requirements of specific areas of science, they lack general applicability. Far from being restricted to medical research, the problem of extrapolation is ubiquitous in science: economists nowadays perform laboratory experiments routinely, for example, yet they are not interested in the behavior of undergraduate students playing the prisoner's dilemma game in a university lab. To generalize the results of such experiments, one also needs information about the way "real people" behave in "real" economic contexts. But if field data are required, what do the experiments contribute in the first place? Without a rigorous analysis and justification, local methods are difficult to improve or extend to other domains. The value of devising a general philosophical account seems non-negligible.

Although a few philosophers have discussed it in the past (LaFollette and Shanks 1995; Guala 1998, 2005; Thagard 1999; Schaffner 2001; Weber 2005), in *Across the Boundaries* Dan Steel (2008) provides the first general account that promises to apply across the natural and social sciences. Steel's theory is based on the method of comparative process tracing—the idea, roughly, that an inference from laboratory conditions to a nonlaboratory (target) system is validated by cross-checking the crucial nodes of the relevant causal mechanisms. I illustrate in more detail the process-tracing account in section 2. In that section I also outline the methodological problems that, according to Steel, the process-tracing approach is meant to successfully tackle. One of the central theses of *Across the Boundaries* is that process tracing succeeds where previous accounts of extrapolation have failed. I am not entirely convinced, and I will suggest that the "analogical" account can provide an adequate account of extrapolation. This is not meant to diminish the value of Steel's contribution—on the contrary, the application of mechanistic concepts and tools marks a crucial step forward in the debate and (as explained in his contribution to this symposium) can be extended to model general forms of analogical reasoning too.

2. The Comparative Process-Tracing Account of Extrapolation. The problem of extrapolation concerns in principle any attempt to generalize from an epistemically privileged system (experimental mechanism, animal model, etc.) to a less privileged target of interest (a nonexperimental situation, human system, etc.). Although such an inference can regard any sort of scientific information, the most common and important instances of extrapolation regard the generalization of causal claims. It is usually our interest in recipes for intervention—and the worry that such interventions may go awfully wrong—that makes extrapolations so important and so delicate at the same time.

In *Across the Boundaries*, Steel (2008) sets out with utmost clarity the

https://doi.org/10.1086/656541 Published online by Cambridge University Press

desiderata that a serious solution to the problem of extrapolation ought to satisfy. Every such account must be able to answer two challenges, the "extrapolator's circle" and the "problem of difference":

> Additional information about the similarity between the model and the target—for instance, that the relevant mechanisms are the same in both—is needed to justify the extrapolation. The extrapolator's circle is the challenge of explaining how we could acquire this additional information, given the limitations of what we can know about the target. In other words, it needs to be explained how we could know that the models and the target are similar in causally relevant respects without already knowing the causal relationship in the target. The second challenge arises from the inevitable presence, in the biological and social sciences, of causally relevant differences between the model and the target. Thus, any adequate account of extrapolation in heterogeneous populations must explain how extrapolation can be possible even when such differences are present. (Steel 2008, 78–79)

Steel's answer to these two challenges is based on a methodology articulated by Darden and Craver (2002) and Glennan (2005) and variously called "forward chaining," "backtracking," or "process tracing." Let us call a set of interconnected, regularly operating causal relationships that generate one or more regularities between (observable or unobservable) events a "mechanism." Clearly much scientific research in the natural and social sciences is devoted to the discovery and theoretical description of such mechanisms. In order to discover the fine-grained structure of a mechanism, one may start from one of its end points (one of the initial causes or one of the final effects) and then step-by-step reconstruct the path that connects it with the other elements of the mechanism via the intermediate nodes. This is, in broad terms, the method of process tracing as applied in the service of mechanism discovery.

Notice that in this case process tracing is used to uncover the structure of a single mechanism. In the case of extrapolation, the scientist is facing a different problem—how to infer from knowledge of a causal mechanism in A the existence of a similar mechanism in B. Comparative process tracing can be articulated in two preliminary steps and an inductive inference:

(1) Learn the mechanism in the model organism [or experimental system], by means of process tracing or other experimental means.
(2) Compare stages of the mechanism in which the two [the experimental and the target systems] are most likely to differ significantly.

https://doi.org/10.1086/656541 Published online by Cambridge University Press

(3) In general, the greater the similarity of configuration and behaviour of entities involved in the mechanism at these key stages, the stronger the basis for the extrapolation. (Steel 2008, 89)

Comparative process tracing is employed primarily to answer the first challenge of extrapolation.[2] The extrapolator's circle is broken by pointing out that comparisons are required only at those nodes where the two mechanisms are most likely to differ. The more that nodes are available for inspection, the stronger the inference is supposed to be (step 3). But full knowledge of the target mechanism is not required in principle. Moreover, if one of the initial causal factors is likely to leave a mark that is transmitted in the causal chain, it will often be sufficient to check downstream differences in the two mechanisms (Steel 2008, 95). Again, full knowledge of the target system is not required for extrapolation to take off.

3. The Analogical Account Revisited. Comparative process tracing is an elegant, general account of extrapolation that promises to be applicable to a wide range of circumstances across the natural and social sciences and that fills a gaping hole in the theory of scientific method. In what follows, however, I want to take issue with Steel's claim that process tracing succeeds where previous accounts have failed. In so doing, I will also address the problem of extrapolation of experimental results in the social sciences, toward which Steel has a somewhat pessimistic attitude. In order to do that, it is necessary to introduce briefly the main competitor to the process-tracing view.

Several philosophers have noticed that extrapolation inferences display many features that are typical of analogical reasoning. Hugh LaFollette and Niall Shanks (1995), for example, argue that inferences from animal models in the biomedical sciences are justified only insofar as the animals are causal analogical models (CAMs) of the target system. The criterion of "CAM-hood" is the absence of relevant causal disanalogies between the model and the target. This criterion violates the requirements that an acceptable account of extrapolation is supposed to satisfy: in order to establish CAM-hood, full knowledge of both target and experimental mechanism is needed, and no extrapolation is possible unless the analogy holds exactly.

This is not surprising—indeed, it is an intended consequence of the CAM account. LaFollette and Shanks have a skeptical agenda and use the idea of CAMs to question the use of animal experimentation in medical

2. Steel (2008) tackles the difference problem by formulating a general principle of "positive causal relevance" that allows the extrapolation of qualitative causal claims. Since this issue is largely independent of the comparative process-tracing account and involves technical issues in probabilistic causal inference, I will not discuss it here.

https://doi.org/10.1086/656541 Published online by Cambridge University Press

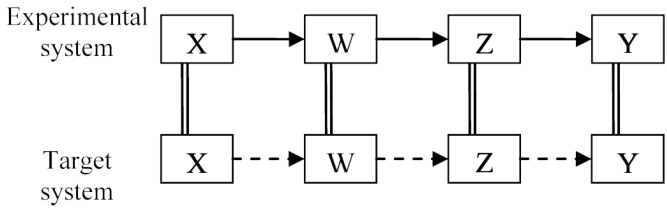


Figure 1. Analogical reasoning and process tracing.

research. But other philosophers have tried to give a positive answer to the challenge of extrapolation. They have tried to show that extrapolation from experimental systems and animal models is possible in principle and that such inferences can in some circumstances be sound. To the extent that such authors have followed LaFollette and Shanks in modeling extrapolation as a form of analogical reasoning, they have the burden of showing that their theories do not lead to the same skeptical conclusions. In *The Methodology of Experimental Economics* and elsewhere (Guala 1998, 2005), I have followed Paul Thagard (1999) in arguing that extrapolation inferences can be reconstructed as analogical inferences of the following kind: (1) humans have symptoms Y, (2) laboratory animals have symptoms Y, (3) in laboratory animals, the symptoms are caused by factor X, (4) the human disease is therefore also caused by X (Guala 2005, 196).

Since analogical inferences are fallible and vary in strength, an analogical account must explicate the conditions that make some extrapolations more reliable than others. Ideally, we would like to show that strong analogical extrapolations are sanctioned by the principles of a general theory of inductive reasoning. At the moment we are still far from achieving this goal, but the process-tracing approach faces the same challenge. According to Steel, the reliability of an inference is directly proportional to the number of stages (in the mechanism connecting X and Y) where comparative checks of dissimilarity can be made and to the reliability of the background knowledge used to identify the stages where the most threatening dissimilarities are likely to be located. This is equivalent to increasing the number of established analogies along the two paths connecting X and Y. In figure 1 the arrows represent causal relations, and the double lines represent analogical relations. The dotted lines stand for the fallible conclusion to be established by extrapolation.

In *Across the Boundaries*, Steel recognizes that the analogical account is compatible with a mechanism-based theory of extrapolation (2008, 151) but claims that my version of the analogical theory is bound to collapse into LaFollette and Shanks's skepticism. Although rhetorically more friendly (because of my repeated suggestions that the problem of extrap-

https://doi.org/10.1086/656541 Published online by Cambridge University Press

olation can and ought to be solved), my analogical account nevertheless would fail to provide an acceptable answer to the challenges of extrapolation. Evidence to back up this charge can be found in those passages where I claim that extrapolation inferences "cannot be strong unless experimental and field evidence have been generated by systems that are similar in *all relevant* respects or, in other words, unless *all* sources of external validity [i.e., extrapolation] error have been taken care of by means of accurate design" (Guala 2005, 199; emphasis added). This makes it sound as if I subscribed to LaFollette and Shanks's criterion of CAM-hood, with the consequence of imposing too strict requirements on reliable extrapolation.

However, part 2 of *The Methodology of Experimental Economics* (Guala 2005) is explicitly concerned with the articulation of a theory of inductive inference that can account for extrapolation. Such a theory "will have at least to be compatible with the theory of internal validity outlined in the first part of the book; but in fact . . . it will turn out that both inferences are instances of the eliminative inductive method" (Guala 2005, 194). The confusion regarding the "relevant similarities" stems from overlooking the theory of induction defended in part 1 of the book.[3] A key idea there is that it is useful to distinguish between objective and subjective notions of support.[4] An objective analysis outlines the conditions for a piece of evidence e to support H independently of what is known or believed about the relation between e, H, and the background conditions K that help us draw the inference from e to H. Its main contribution is normative—it allows the identification of what ought to be done and ought to be known in order to design a good experiment or to make a strong inductive inference. But in judging whether an inference is warranted or not, a subjective element is necessarily involved: in practice, one can only check the dissimilarities that are deemed relevant by one's current background knowledge.

The key word then is "relevant": interpreted objectively, the requirement that the experimental and target systems ought to be similar in all causally relevant respects would indeed collapse my analogical account onto LaFollette and Shanks's CAM theory. But the analogical account, in fact, only requires that we check all the (dis)similarities that are deemed relevant by our current background knowledge. So it is not necessary that the two causal mechanisms are similar at each and every stage—they must only

3. For more details, see Guala (2005, chap. 6; 2008). The theory of induction that I defend belongs to the "testing approach" to scientific inference that has been variously elaborated by Giere (1983), Mayo (1996, 2005), and others.

4. Guala (2005, 115). The distinction is far from new: similar ideas have been proposed before by, e.g., Peter Lipton and Peter Achinstein.

https://doi.org/10.1086/656541 Published online by Cambridge University Press

be similar to the extent that it is supposed to matter, according to our current best scientific knowledge.

Empirical studies of scientific practice, in fact, reveal that experimental results are accepted by the scientific community when experimenters have shown that all the factors that may lead to error have been controlled for. But the universal quantifier should not be read metaphysically as "all possible sources of errors"—for some of them may not be even known—but rather as "all the errors that the community of scientists working in that area consider worth checking." The list is necessarily going to be finite and to an extent will be determined by conventional and sociological factors (as well as by past experience, of course; see Guala 2005, 130–32). In the case of extrapolation, errors are possible at every node of the causal processes. But obviously we do not want to check all the nodes in the target mechanism—if we could do that, we would not need laboratory experiments in the first place.

4. An Example. Overall then, there is no difference in spirit between the analogical and the comparative process-tracing account. The latter's contribution consists primarily in the detailed articulation of the methodological strategies adopted by scientists to tackle extrapolation problems (and their rationale). It may be useful to examine to what extent comparative process tracing departs from the analysis provided by supporters of the analogical approach.

The central instance of analogical reasoning in *The Methodology of Experimental Economics* is the case of the outer continental shelf (OCS) auctions. These auctions are used routinely to sell oil leases in the Gulf of Mexico, using a first-price, sealed-bid mechanism. In the 1970s field data suggested the existence of systematic overpricing in these auctions, resulting in lower-than-average profits (and sometimes losses) for oil companies. Some economists suggested that the auction rules may be partly responsible for this: in a first-price, sealed-bid auction, each participant must consider that all estimates of the real (but unknown) value of a lease are likely to be inaccurate to some extent. If everyone bid their own estimate, then, the lease would go to the bidder who overestimates that value the most. If winners are systematically "cursed" this way, the average profits will be lower than one would expect in an efficient market.[5]

Important evidence in favor of the winner's curse hypothesis came from experiments run by John Kagel and Dan Levin (1986). The crucial advantage of the experimental method is that the true value of an auctioned item is known in the lab, and bidders' estimates can be controlled to a

5. See Thaler (1988) for an introductory account. Guala (2005, chap. 9) includes a detailed methodological analysis.

https://doi.org/10.1086/656541 Published online by Cambridge University Press

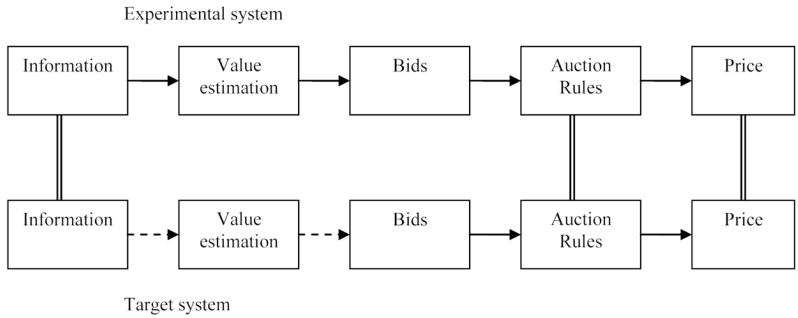


Figure 2. Analogical reasoning in the OCS example.

greater extent than in the field. Kagel and Levin provided their subjects with some information regarding the value of the auctioned item, drawn from a random distribution centered on its true value. They then observed whether bidders were able to revise their bids so as to take into account the possibility of overbidding and thus avoid the winner's curse effect.

The winner's curse was observed in these experimental auctions. Kagel and Levin, however, were also concerned with demonstrating that the laboratory result could be extrapolated to explain the target phenomenon (i.e., real bidding in the OCS auctions). They showed that variations in the amount and quality of information provided to bidders in the laboratory were correlated with variations in prices that mirror those observed in the field when equivalent variations in information occur (Guala 2005, 192–94). One important point is that this analogy was not secured by establishing a point-by-point similarity between all the causal factors that matter. This is not what the analogical account requires. The analogy was backed up by looking for consequences or side effects of the (relevantly similar) mechanisms that supposedly govern the laboratory and the target systems. In process-tracing terminology, this is equivalent to establishing that there is a relevant similarity at a "downstream" stage of the mechanisms and using this evidence to back up the claim that the mechanisms are similar enough in the upstream part too.

Figure 2 highlights where the analogical correspondences take place. As before, the arrows represent causal relations and the double lines analogical correspondences. The mechanisms are simplified for representational ease (e.g., there should be a box representing bidders' preferences at the beginning, which I have omitted). Notice that in the laboratory some analogical correspondences are constructed by design: the auction rules in the experiment, for example, are designed so as to resemble as closely as possible those of the target. The information provided to sub-

https://doi.org/10.1086/656541 Published online by Cambridge University Press

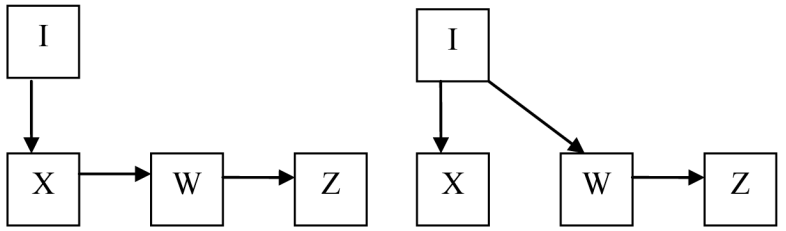


Figure 3. Structure-preserving versus structure-altering intervention.

jects is also meant to mimic the characteristics of the information that is available in the field. Checking that no dissimilarities occur at each and every stage of the mechanisms would not be feasible, though—it is precisely because some factors (bidders' estimations and bidders' computational abilities, for example) are not observable in the field that experimental investigation is extremely valuable. But by comparing analogous variations in information and prices at a downstream stage of each mechanism, it is possible to extrapolate the winner's curse result from the laboratory to the field.

The analogical approach, to sum up, addresses the extrapolator's circle in a satisfactory manner, although this is obscured by Steel's choice of examples. Of the three main examples discussed in Guala (2005), Steel ignores precisely the OCS case, the one that is central in the presentation of the analogical approach. This does not mean that the comparative process-tracing account is wrong—on the contrary, it is largely correct, and it brings the discussion to a level of clarity that is lacking in earlier accounts. The strategy of checking downstream stages of the mechanisms, in particular, is a novel contribution that was made possible by framing the issue in terms of comparative process tracing. Moreover, process tracing is important for Steel's analysis of the special difficulties facing extrapolation in the social sciences. I turn to this aspect of the problem in the next section.

5. Extrapolation in the Social Sciences. Steel (2008, chap. 8) highlights two problems that often affect attempts at extrapolating in the social sciences. One is the problem of structure-altering intervention; the other is the problem of uncertainty about mechanisms. An intervention is "structure altering" if it changes not only the value of an experimental variable (the treatment) but also other features of the causal structure that the experimenter is investigating. Figure 3 includes examples of a proper, "structure-preserving" and an improper, "structure-altering" intervention. Suppose that before the intervention, X causes W, which in turn causes Z ($X \rightarrow W \rightarrow Z$). A proper intervention (on the left) is "surgical" (Wood-

https://doi.org/10.1086/656541 Published online by Cambridge University Press

ward 2003) in the sense that it cuts all and only the causal relations that determined the value of X in the original structure. A structure-altering intervention (on the right), in contrast, changes both the value of X and the value of W.

Steel uses the concept of structure-altering intervention mainly to highlight the difficulties of extrapolating in some social science contexts (he is particularly interested in a field experiment of welfare reform). He argues that because there is a very high chance that any policy intervention is going to be structure altering, when implemented in the field, knowledge about the causal structure collected in experiments will be of very limited use. Consider the following analogy: it is quite useful to know that setting the timer of my clock at 8:00 a.m. will cause it to ring in time for me to go to work. But if, when I actually manipulate the timer, I also switch the clock off, this knowledge will be of no use because it will not give me an adequate recipe for intervention.

However, it seems here that we are trying to draw an inference from the wrong kind of experiment. Surely what we should do is design an experiment where the manipulation is structure altering in the same way as the policy intervention we are interested in. There may be big, practical hurdles that make the design of such experiments difficult or impossible, of course. But in other cases the problem may be solved and extrapolation inferences drawn that teach us something about social policy and intervention.

The second problem highlighted by Steel concerns uncertainty about the causal mechanisms. Steel illustrates the problem with reference to the case of preference reversals (PRs) experiments. Preference reversals are among the most widely discussed phenomena in experimental economics and psychology (see Thaler and Tversky [1990] for a survey). Subjects in a PR experiment are asked to price a series of bets with specific characteristics and then to choose among pairs of the same bets. A remarkable number of subjects choose the bet, in such pairs, that had been assigned the lower price. Assuming a principle of "procedural invariance"—that pricing and choosing are equivalent ways of manifesting one's preferences—this behavior is a clear symptom of preference inconsistency.

Steel (2008, 169–73) highlights economists' and psychologists' disagreement about the causes of PRs as the main obstacle toward the application of the comparative process-tracing method. To an extent, he is right. But this may reflect simply a particular stage of research on PRs—there is nothing, in principle, that seems to prevent behavioral scientists from reaching an agreement on the relevant causal mechanism. Indeed, it would be easy to cite other examples of experimental phenomena whose mechanisms are satisfactorily understood and where disagreement remains con-

https://doi.org/10.1086/656541 Published online by Cambridge University Press

cerning only marginal issues (the contribution to public goods may be a case in point—see, e.g., Guala [2005], chap. 2).

The real problem with the extrapolation of PRs, in my view, is rather different. Debates about the scope of PRs (or their extrapolability) are ultimately invalidated by lack of clarity concerning the target system. Critics of PR experiments often focus on the chance of observing PRs in ideal competitive markets but fail to study real economic settings in which PRs are likely to manifest (Guala 2005, 225–29). But until a target has been identified, the method of comparative process tracing (or analogical inference, for that matter) cannot even take off.

6. Extrapolation without Process Tracing. I want to finish on a rather speculative note. The comparative process-tracing approach highlights the importance of articulating causal mechanisms for extrapolation. While this seems right in a number of cases, the full reconstruction of the experimental mechanism may be unnecessary in others. Consider the OCS case again. As a matter of fact, it would have been more realistic to write dotted arrows linking information, value estimation, and bids in the experimental system represented on top in figure 2. The detailed mechanism behind the winner's curse phenomenon is largely unknown and, somewhat surprisingly, not even studied. The reason is simply that such knowledge is required neither for internal nor for external validity purposes. All that is required is a definition of the winner's curse as a deviation from perfect rationality.

Under the perfect rationality hypothesis that is commonly made in economic models, each bidder should revise her value estimate (x_0) of the auctioned bid as if the estimate she had formed on the basis of her own information (X_i) were the highest estimate (x_1) among those of all market competitors. The winner's curse can be defined simply as a failure to revise one's estimate conditional on this assumption and, in particular, as an overestimation of the following kind:

$$E[x_0|X_i] > E[x_0|X_i = x_1].$$

This clearly falls short of a proper articulation of the mechanism (or mechanisms) governing the winner's curse phenomenon, even in the lab.[6] And yet, it is perfectly adequate for extrapolation purposes. Large parts of the mechanism can be "black boxed" as long as there are good reasons to believe that they are analogously instantiated in the laboratory and the target system. So here is an example where an analogy is drawn to

6. Another reason why experimental economists and psychologists are often more interested in phenomena than in mechanisms is that such phenomena often result from a combination of different mechanisms "pushing" in the same direction, so to speak (see Sugden 2005).

https://doi.org/10.1086/656541 Published online by Cambridge University Press

support extrapolation (value estimation in the experiment is similar to that in the target) while lacking a detailed specification of the causal mechanism both in the target and in the experiment. Steel's example (in his contribution to this symposium) of establishing an analogy between two engines on the basis of their sound—rather than detailed analysis of their causal mechanisms—is another case in point.

Perhaps when we have a general theory of inductive inference that can be applied to extrapolation, we will discover that analogical reasoning is more fundamental than mechanistic reasoning. Steel's paper moves one step in this direction, I believe. Surely recent developments in the philosophy of causality suggest that the notion of causation that is so important for science (and experimental science, in particular) may be too rich to be analyzed using mechanistic concepts only.[7] Mechanistic accounts typically fail to capture the "difference-making" sense that is embedded in many causal expressions that we use in everyday and scientific discourse (e.g., John's failure to stop the murderer caused Jim's death, or, in our case, the bidders' failure to revise their estimates causes the winner's curse phenomenon). Because the difference-making aspect of causality is central in experimental discourse, we need more general methods of extrapolation that are able to account for it.

REFERENCES

Cartwright, Nancy. 2007. *Hunting Causes and Using Them*. Cambridge: Cambridge University Press.

Darden, Lindley, and Carl Craver. 2002. "Strategies in the Interfield Discovery of the Mechanism of Protein Synthesis." *Studies in History and Philosophy of Biological and Biomedical Sciences* 33:1–28.

Giere, Ronald. 1983. "Testing Theoretical Hypotheses." In *Testing Scientific Theories*, ed. John Earman, 269–98. Minneapolis: University of Minnesota Press.

Glennan, Stuart. 2005. "Modeling Mechanisms." *Studies in History and Philosophy of Biological and Biomedical Sciences* 36:443–64.

Godfrey-Smith, Peter. 2009. "Causal Pluralism." In *Oxford Handbook of Causation*, ed. Helen Bebee, Peter Menzies, and Christopher Hitchcock. Oxford: Oxford University Press.

Guala, Francesco. 1998. "Experiments as Mediators in the Non-laboratory Sciences." *Philosophica* 62:901–18.

———. 2005. *The Methodology of Experimental Economics*. New York: Cambridge University Press.

———. 2008. "The Experimental Philosophy of Experimental Economics: Replies to Alexandrova, Hargreaves-Heap, Hausman and Hindriks." *Journal of Economic Methodology* 15:224–31.

Kagel, John, and Dan Levin. 1986. "The Winner's Curse Phenomenon and Public Information in Common Value Auctions." *American Economic Review* 76:894–920.

LaFollette, Hugh, and Niall Shanks. 1995. "Two Models of Models in Biomedical Research." *Philosophical Quarterly* 45:141–60.

7. See, e.g., Cartwright (2007), Reiss (2008), and Godfrey-Smith (2009).

https://doi.org/10.1086/656541 Published online by Cambridge University Press

Mayo, Deborah. 1996. *Error and the Growth of Experimental Knowledge*. Chicago: University of Chicago Press.
———. 2005. "Evidence as Passing a Severe Test: Highly Probable versus Highly Probed Hypotheses." In *Scientific Evidence: Philosophical Theories and Applications*, ed. Peter Achinstein. Baltimore: Johns Hopkins University Press.
Reiss, Julian. 2008. "Causation in the Social Sciences: An Inferentialist Perspective." Paper presented at the 2008 International Philosophy of Science conference, Dubrovnik.
Schaffner, Kenneth. 2001. "Extrapolation from Animal Models: Social Life, Sex, and Super Models." In *Theory and Method in the Neurosciences*, ed. Peter Machamer, Rick Grush, and Peter McLaughlin. Pittsburgh: University of Pittsburgh Press.
Steel, Daniel. 2008. *Across the Boundaries: Extrapolation in Biology and Social Science*. New York: Oxford University Press.
Sugden, Robert. 2005. "Experiments as Exhibits and Experiments as Tests." *Journal of Economic Methodology* 12:291–302.
Thagard, Paul. 1999. *How Scientists Explain Disease*. Princeton, NJ: Princeton University Press.
Thaler, Richard H. 1988. "Anomalies: The Winner's Curse." *Journal of Economic Perspectives* 2:191–202.
Thaler, Richard H., and Amos Tversky. 1990. "Anomalies: Preference Reversals." *Journal of Economic Perspectives* 4:201–11.
Weber, Marcel. 2005. *Philosophy of Experimental Biology*. Cambridge: Cambridge University Press.
Woodward, Jim. 2003. *Making Things Happen*. New York: Oxford University Press.

https://doi.org/10.1086/656541 Published online by Cambridge University Press