

REPRINT

DEVELOPMENTS
IN THE METHODOLOGY OF
SOCIAL SCIENCE

Edited by

WERNER LEINFELLNER AND ECKEHART KÖHLER

University of Nebraska, Department of Philosophy



D. REIDEL PUBLISHING COMPANY

DORDRECHT-HOLLAND / BOSTON - U.S.A.

BERNARD GROFMAN

RATIONAL CHOICE MODELS AND
SELF-FULFILLING AND
SELF-DEFEATING PROPHECIES

ABSTRACT. If individuals act so as to maximize expected utility then both self-fulfilling and self-defeating prophecies may be integrated into a rational choice model in such a way as to maintain predictability (in probabilistic terms) of individual actions.

The self-fulfilling (or self-defeating) prophecy, has stalked the pages of anti-behavioralist tracts ever since its introduction into the social sciences by a social scientist of impeccably behaviorist credentials, Robert Merton.¹ A self-fulfilling (self-defeating) prophecy consists of "predictions that are false (true) to the actual facts at the time the predictions are made, but that nonetheless turn out to be true (false) because of the actions taken as a consequence of belief in the predictions."² The first point we wish to make about self-fulfilling (self-defeating) prophecies is that it is certainly true that actors' beliefs can lead to actions which, intentionally and/or unanticipatedly verify or falsify those very beliefs. For example, although the U.S. Bank (a private bank in New York City, despite its name) was in no trouble, many people thought it was in dire trouble and would soon fail. This belief led to a run on the bank, so that the organization was, in fact, compelled to go into bankruptcy.³ But what we may ask are the implications of the existence of self-fulfilling and self-defeating prophecies for the logic of social science inquiry?

Self-fulfilling and self-defeating prophecies both involve actions which have unanticipated consequences. That these consequences may verify or falsify previously held beliefs is really accidental, by which we mean no more than that there may be unanticipated consequences without either a self-fulfilling or self-defeating prophecy being involved. Does the existence of the possibility of a self-fulfilling prophecy or more generally, of consequences unanticipated by the actor himself, necessarily make prediction of an actor's actions impossible? Consider what would happen if an actor were informed that his withdrawal of funds from a bank might trigger a run on the bank? If he believed that the bank would fail anyway (a necessary antecedent condition for the existence of a self-fulfilling

prophecy), or if he believed that the bank would not fail regardless (a necessary condition for the existence of a self-defeating prophecy), he would in either case, we could hypothesize, be likely to withdraw funds, regardless of the possible consequences for the bank, whenever his expected utility from withdrawal of funds =

$$\begin{aligned} & \text{prob (bank failure/withdrawal)} \\ & \times U \text{ (bank failure/withdrawal)} \\ + & \text{prob (no bank failure/withdrawal)} \\ & \times U \text{ (withdrawal/no bank failure)} \\ + & \text{prob (bank failure/withdrawal)} \\ & \times U \text{ (withdrawal/bank failure)} \end{aligned}$$

was greater than his expected utility from not withdrawing funds.

$$\begin{aligned} = & \text{prob (bank failure/no withdrawal)} \\ & \times U \text{ (bank failure/no withdrawal)} \\ + & \text{prob (no bank failure/no withdrawal)} \\ & \times U \text{ (no withdrawal/no bank failure)} \\ + & \text{prob (bank failure/no withdrawal)} \\ & \times U \text{ (no withdrawal/bank failure)} \end{aligned}$$

There is nothing *in principle* which prevents us from ascertaining these subjective probability and utility assessments and predicting the actor's behavior accordingly. Of course, it is true that making such predictions will require knowing the actor's information and conditional assessments about states of the world, but this is no inherent bar to the possibility of prediction, but simply a specification of facts which a prediction/explanation of the actor's behavior will need to incorporate, if the generalizations to be used specify information and expectations as relevant variables.⁴ More generally, "if the knowledge that men possess of social processes is a variable that enters into the determination of social phenomena, there are no a priori grounds for maintaining that changes in that variable and the effects they may produce cannot be the subject of social laws."⁵

Bandwagon and underdog effects of election predictions are respectively, species of self-fulfilling and self-defeating prophecies. Herbert Simon has shown that, given assumptions about the strength of these effects in the voting population, a pollster can, in principle, announce his prediction in such a form that, despite the changes in voters' electoral behavior as a

result of the information provided by the predictions, the vote distribution in the election will remain unchanged.⁶ Relatedly, John von Neumann has shown that, given certain plausible assumptions, in many kinds of games there will exist strategies such that a player will not deviate from them even if he somehow learns in detail his opponents' plans.⁷ More generally, the effects over time of new knowledge in behavior may be dealt with in terms of feedback loops and the problem of predictability posed in terms of stable solutions to sets of recursive equations – problems which are in form identical to those dealt with in the theory of automata; demonstrating a basic similarity in behavior systems whether human, organic, or what have we.⁸

*Department of Political Science,
State University of New York at
Stony Brook*

NOTES

¹ See David Easton, *The Political System*, Knopf, New York, 1953, pp. 24–31.

² *Social Theory and Social Structure*, rev. ed., The Free Press, New York, 1957, Chapter 11.

³ Nagel, *The Structure of Science* p. 469, adapted from Merton, *Ibid*, Chapter 11.

⁴ Cf. Easton, *op. cit.*

⁵ Nagel, *op. cit.*, p. 471. It is important in this connection to recall that laws, whether in the social or the natural sciences, are conditional in form. That the antecedent conditions for a law may not be obtained does not invalidate the law. "(T)he fact that men generally avoid exposure to the fumes of hydrocyanic acid when they become familiar with the law that if the gas is inhaled, death rapidly follows is (not) a disproof of this law, and might, indeed, suggest the probabilistic 'law' that 'Men generally avoid exposure to the fumes of hydrocyanic acid when they become familiar with the law that if the gas is inhaled death rapidly follows.'" (Nagel, *Ibid*, p. 471).

⁶ Herbert Simon, 'Bandwagon and Underdog Effects of Electoral Predictions', in *Models of Man*, John Wiley and Sons, New York, 1957, Chapter 5.

⁷ John Von Neumann and Oskar Morgenstern, *The Theory of Games and Economic Behavior*, Princeton University Press, 1947. For a more elementary and quite delightful introduction to game theory see J. J. McKinsey, *The Compleat Strategist*, McGraw-Hill, New York.

⁸ We might also note that the prediction of what new knowledge will be achieved should not be confused with the knowledge itself. One does not need to be a geneticist to predict that certain breakthroughs in genetics will be achieved, or indeed, to predict some likely (or at least possible) social consequences of these breakthroughs. This distinction is completely missed by Mulford Sibley when he asserts that we cannot foretell what new knowledge will be discovered in the future, for if we could, we ourselves would be the discoverers. (Sibley, 'The Limitations of Behavioralism', in J. C. Charlesworth (ed.) *Contemporary Political Analysis*, The Free Press, New York, 1967, p. 65.)