

## When Analytic Narratives Explain

**Anna Alexandrova**

*University of Missouri St Louis*

*Department of Philosophy*

*alexandrova@umsl.edu*

---

### Abstract

Rational choice modeling originating in economics is sweeping across many areas of social science. This paper examines a popular methodological proposal for integrating formal models from game theory with more traditional narrative explanations of historical phenomena, known as “analytic narratives”. Under what conditions are we justified in thinking that an analytic narrative provides a good explanation? In this paper I criticize the existing criteria and provide a set of my own. Along the way, I address the critique of analytic narratives by Jon Elster.

### Keywords

analytic narratives, rational choice models, causal explanation, de-idealization

### 1. Introduction<sup>1</sup>

Rational choice modeling originating in economics is sweeping across many areas of social science. This paper examines a popular methodological proposal for integrating formal models from game theory with more traditional narrative explanations of historical phenomena. This proposal is known as “analytic narratives”. Although the idea of coupling models with narratives is not new,<sup>2</sup> an explicit discussion of its merits and disadvantages

---

<sup>1</sup>) This paper benefited from comments I received from the participants of Washington University in St Louis Political Theory Workshop, Robert Northcott and an anonymous referee. Nancy Cartwright helped me to develop the Open Formulae approach in my PhD dissertation at UC San Diego.

<sup>2</sup>) For a history of this approach see M. S. Morgan, “The curious case of the Prisoner’s Dilemma: Model Situation? Exemplary narrative?” in A. Creager, M. Norton Wise and

starts with a 1998 book *Analytic Narratives* edited by Robert Bates, Avner Grief, Margaret Levi, Jean-Laurent Rosenthal and Barry Weingast. This book contains five chapters in which each contributor proposes an analytic narrative of some puzzling historical phenomenon (the stability of peace between clans of 12th century Genoa, the institutional foundation of US federalism, the rise and fall of the International Coffee organization, etc.), and an introduction in which the authors together explore broad methodological rules for building an analytic narrative.

Roughly, the proposal is as follows: learn about the phenomenon of interest (either by studying the past directly or reading the already existing historical accounts) so as to be able to mentally isolate a rational strategic interaction between actors, which can then be formalized in a rational choice model. Game theory models in extensive form seem to be the most popular, though purportedly do not have to be. This model should bear some degree of similarity to the choices and trade offs the actors faced in the phenomenon in question. The model can then be supplemented with a narrative, and/or a Geertzian “thick description” – i.e. an explanation of the meaning actors attach to their actions, circumstances and surroundings, that is their significance within the local culture. The rational choice model may predict an outcome that fails to occur or not predict a specific outcome at all (in case of multiple equilibria), in which case the narrative will be engaged to explain what other factors, which the model leaves out, account for the outcome.

Its advocates claim that this methodology can unify nomothetic and ideographic approaches to social science. The narrative captures the uniqueness of the situation under investigation, while the model captures the general features of the type of phenomena under which this situation falls. Whether or not this is true, here I shall assume that at least one legitimate way of explaining social phenomena is to offer a causal explanation. This is precisely the sort of explanations that the proponents of ANs aim for. A causal explanation does not need to be a law-based explanation and it does not need to be deductive as the covering law model of explanation assumed.<sup>3</sup> There are a number of ways to flesh out what a causal explanation amounts

---

E. Lünebeck *Science without Laws* (Durham: Duke University Press, 2005). For the role of narratives in model-based explanation see M. S. Morgan, “Models, Stories, and the Economic World” *Journal of Economic Methodology* 8/3 (2001), 361–84.

<sup>3</sup> C. Hempel, *Aspects of scientific explanation and other essays in the philosophy of science* (New York: Free Press, 1965).

to. Some philosophers make use of the notion of mechanism,<sup>4</sup> others analyze causal explanation by using the notion of ideal intervention,<sup>5</sup> etc. For the purposes of this paper no particular account of causal explanation, or causation, needs to be presupposed. I simply assume that one way to explain a social phenomenon is to state how some combination of causes brought it about.

What makes for a good analytic narrative (hereafter AN)? Under what conditions are we justified in thinking that an AN provides a good causal explanation? First a note of caution. On the AN approach the explanans is a complex whole – it is a formal model embedded in a narrative. So to specify conditions for a successful AN we need to specify the role of the model as well as the role of the narrative in an explanation of a phenomenon. In this paper I am only concerned with the role of the model in explanation.<sup>6</sup> Figuring out just what this role is is urgent both for social science and for philosophy. It is urgent for social science because we need to know whether to train our students to write analytic narratives.

And it is urgent for philosophy because it taps in into a classic and still widely discussed issue of the role of idealizations in science. It is no news that the sort of models around which ANs are built are strongly idealized. They make strict assumptions about who the players are, which options are open to them, what they assign value to, what information they have access to, and perhaps, most controversially, how they reason. The rationality assumed in the vast majority of these models is that of expected utility maximization. This assumption is used in various sorts of equilibrium solutions. For example, Bayesian Nash equilibrium assumes that players pick the best action relative to their beliefs and to what they expect other players to do. Further unrealistic assumptions include mathematical features of players' utility functions (for example, their differentiability) and of probability distributions (such as their continuity, uniformity, etc). That

---

<sup>4</sup> M. Bunge, "Mechanism and Explanation", *Philosophy of the Social Sciences*, 27 (1997) 4: 410–465; J. Elster, *Nuts and Bolts for the Social Sciences* (Cambridge, UK: Cambridge University Press, 1989); P. Hedström, *Dissecting the Social: On the Principles of Analytic Sociology* (Cambridge University Press, 2005); P. Hedström and P. Swedberg, *Social Mechanisms: An Analytical Approach to Social Theory* (Cambridge University Press, 1998).

<sup>5</sup> J. Woodward, *Making Things Happen: A Theory of Causal Explanation*. (Oxford: Oxford University Press 2003).

<sup>6</sup> See P. A. Roth, "How Narratives Explain" *Social Research* 56/2 (1989), 449–478, among others on the role of narratives in explanation.

these assumptions more often than not fail to hold in the real world situations these models are applied to is no secret. How then do these models nevertheless explain? An account that spells out conditions for a successful AN must, among other things, answer this question.

In this paper, I discuss the existing accounts of ANs by their founders and a critique by Jon Elster. I then put forward my own criteria for successful explanation by AN. The upshot is that idealizations in a model do not preclude this model's participation in explanation. However, since rational choice modelers are by and large unable to demonstrate that these idealizations do not matter, the models cannot be treated as specifying the causal mechanism of the situation to which they are applied. In the best case, the model can provide categories in terms of which a hypothesis about this causal mechanism can be constructed. But even then the model cannot be said to explain the phenomenon or to get confirmed by it. Moreover, models can and do fail to provide the right categories for constructing a hypothesis about the mechanism in question. I thus urge practitioners of AN not to overstate the importance of rational choice models in historical explanation.

## 2. An Example of an Analytic Narrative

Let us start with what I take to be an example of a successful analytic narrative. Although its author does not identify it as such, political scientist Kenneth Schultz's explanation of how domestic political competition affected resolution of the Fashoda crisis of 1898 is plausibly an analytic narrative. There is a game theoretical model called the Bargaining Game With Strategic Opposition, plus a detailed narrative of the case. The mechanism suggested by the model is claimed to explain the outcome of the crisis.

The Bargaining Game with Strategic Opposition developed by Kenneth Schultz in his 2001 book *Democracy and Coercive Diplomacy*<sup>7</sup> is intended to explain the effect of domestic politics in democracies on their ability to bargain with other states. The game has three actors: government, its inter-

---

<sup>7</sup> K. Schultz, *Democracy and Coercive Diplomacy* (Cambridge: Cambridge University Press, 2001).

nal opposition and a target state. The government is in competition with the target state over some good, and in competition with the opposition for public office. The government has full control over foreign policy decisions, and the opposition has access to information about the expected value of going to war with the target. Because of the relative openness of public debate in a democracy, the target state has access to the opposition's official opinion on whether or not the government should go to war. The good in dispute is divisible and the actors are either risk-neutral or risk-averse. Government and opposition act so as to maximize their probability of (re)election. The target is assumed to value the good for what it is. The electorate in the challenging country values the government's ability to secure more rather than less of the good and at lower rather than higher price. The opposition's performance is judged on its support of high-return low-cost wars or criticism of high-cost low-return wars. These are the main assumptions of the game.

The government's strategy is a function of the expected value of war. Depending on this, the government will issue a genuine challenge or bluff. Opposition will seek to support use of force by the government if the expected payoff of war is high and hence will only support challenges whose probable outcome it considers attractive. Since the opposition has no motivation to collude with the government in a bluff, the target will be able to know the real costs of the potential conflict to the government and thus to judge the seriousness of the government intentions. If the opposition is rational and office seeking, a supported threat is more credible, as far as the target is concerned, than an opposed one. A credible threat increases the probability of the target conceding without a militarised conflict. A threat with which the opposition would be likely to disagree would carry less credibility and is more likely to be interpreted as a bluff by the target. Thus, democracies have fewer opportunities to exploit private information. The presence of the opposition party has a dual effect on the use of coercive diplomacy by democracies: *democracies threaten more selectively than non-democracies but when they do the threats are more effective* (for inducing concessions on the part of the target without fighting a war).

On the basis of this model Schultz seeks to explain the outcome of the Fashoda crisis in 1898. In this crisis the British issued a threat to challenge the French incursion into colonial Egypt. The threat was supported by the opposition party and was effective enough to cause the French to back

away. The model is also claimed to explain why Hitler was confident enough to remilitarize the Rhineland shortly before WWII (because the threat of the allied governments was not credible enough). All these situations, Schultz argues, share the sort of trade-offs that the participants in an international bargaining situation come up against when they are deliberating on a course of action, and these trade-offs are represented by the model. He marshals impressive historical evidence from both the British and French sources that the opposition party's support of the Prime Minister's threat made this threat more credible. The French gave up because they were convinced that if they stood firm the British would carry out their threat and send troops to retain that part of their colony. And we see that the British government knew that the opposition party would support their threat and went ahead with the threat for this reason.

### 3. Evaluating Analytic Narratives

What makes this AN a successful explanation of the Fashoda crisis? In order to answer this question we need to get clear on the role of the bargaining model in Schultz's explanation. What criteria must be satisfied for this model to play an appropriate role in explanation?

I shall answer this question by discussing the criteria for evaluating ANs proposed by the five original authors and recently further elaborated by Margaret Levi.<sup>8</sup> In each case I shall offer reasons for or against the criterion and propose new versions where needed. The proposed criteria are as follows:

#### 3.1 *Logic of the Model*

Conclusions must be precise and follow deductively. That is, it is a matter of logic that given the assumptions of the Bargaining Model, it **must** be the case that the presence of the opposition party causes governments to be more careful about issuing threats and makes these threats more effective when they are issued.

Levi and her co-authors qualify this criterion by noting that models often do not yield a unique equilibrium, that is they do not make a unique

---

<sup>8</sup> M. Levi, "Modeling Complex Historical Processes with Analytic Narratives" in R. Mayntz (ed.) *Akteure, Mechanismen, Modelle: Zur Theoriefähigkeit makrosozialer Analysen*, 108–127 (Frankfurt/Main: Campus Verlag, 2002).

prediction about what the actors will do in the situation facing them. This is known as the problem of equilibrium selection. In these cases the co-authors allow the narrative to step in. If the model does not predict a precise outcome, then some contextual factor in the narrative might help us to explain why a certain equilibrium path was chosen and not another. However, they also insist that the model must contain the “essential” forces responsible for the phenomenon, so that one does not have to appeal “too often” beyond the model to the narrative.<sup>9</sup>

So a better formulation of this criterion for Levi and her co-authors is as follows: model’s predictions must follow as a matter of logic, but if no precise prediction follows, the narrative may be used as a supplement, so long as the model is the “primary” source of explanation.

What is the motivation behind this qualification and its subsequent restriction? On the face of it, it is bizarre to first combine models and narratives, but then insist that the model should do the lion share of explanation. One possibility is that with this requirement the AN proponents hope to prevent “curve fitting”. If the narrative is always invoked to correct for the model’s lack of correct prediction, then the model is treated as able to accommodate any fact at all, and no fact is incompatible with it.

I doubt that involving the narrative in the explanation has the effect of making ANs unfalsifiable. Why this is so will become clearer later in this article, but a brief explanation is in order now. As deductively closed systems in which assumptions and derivations are connected by logic, models of game theory are not, strictly speaking, falsifiable. What is falsifiable is a hypothesis that researchers make with the model’s help about the phenomenon to which they apply the model. But this hypothesis is not identical with the model’s prediction! As will become clear in the next section, the model only informs, but does not determine, the hypothesis. So the criterion in question amounts to requiring that the model, rather than the narrative, be primarily responsible for the hypothesis. The injunction seems to be: avoid using the narrative to formulate the hypothesis that is supposed to explain the phenomenon featured in this narrative. Curve-fitting, in this sense, amounts to constructing the explanation of the phenomenon on the basis of the data presented in the narrative, rather than from other sources, say, more general theories of institutions. Clearly, this does not make explanations unfalsifiable.

---

<sup>9</sup> Levi “Modeling Historical Processes”, 13.

But is avoiding curve fitting in this sense a good idea nevertheless? In my view, whether or not an AN is curve-fitted in the above sense, so to speak, does not affect its explanatory status.<sup>10</sup> An AN can offer a good explanation of a singular phenomenon even if the narrative, more than the model or some general theory, is responsible for this explanation. A successful explanation, in the sense discussed here, requires a well confirmed causal claim about why and how a certain outcome obtained in the historical case in question. Whether this claim is generated by fitting a model to the phenomenon or by other methods is irrelevant.

Of course, successful explanations are not the only goal of science. We may also want explanations that are generalizable, simple, elegant, fit into a particular theoretical framework, etc. But these are more plausibly requirements of a progressive research program, and thus distinct from requirements for a good explanation.<sup>11</sup> Whether ANs is a progressive research program is a complex question that requires weighing of many different requirements against each other, no doubt with a heavy dose of pragmatic considerations. Perhaps something interesting and informative can be said on this matter, but I will not attempt that.

So I leave the first requirement as is: correct derivations from the model matter indeed, but sometimes, as in the case of multiple equilibria, they are not sufficient. I do not favor, in addition to that, placing restrictions on just how much of the explanation should be informed by the model rather than the narrative. Far more interesting for our purposes is the second proposed requirement on ANs.

---

<sup>10</sup> There are, of course, other senses of curve fitting which are more problematic than this. Curve fitting that is typically criticized by philosophers of science refers to ad hoc modification of hypotheses just to accommodate the data without concern for independent plausibility of this hypothesis (for a recent discussion see E. Sober and C. Hitchcock, “Prediction Versus Accommodation and the Risk of Overfitting”. *British Journal for the Philosophy of Science* 55, (2004), 1–34). I do not believe that this is the sense of curve fitting in question here, since independent plausibility of the hypothesis gets taken into account when this hypothesis is constructed on the basis of the model and the narrative.

<sup>11</sup> One might object that simple historical explanations are better than the complex ones. In my view simple explanations are only better if they are actually better supported by empirical evidence. Indeed most philosophers of science consider simplicity to be a theoretical virtue that only enters the picture when we have two explanations that are equally well supported by empirical evidence (see A. Baker, “Simplicity”, *Stanford Encyclopedia of Philosophy*, ed. E. Zalta, (2004) URL: <http://plato.stanford.edu/entries/simplicity/#4>).

### 3.2 *Assumptions of the Model must “fit the facts”*<sup>12</sup>

In imposing this requirement the authors wish to explicitly reject economist Milton Friedman’s instrumentalism about models. In his classic 1953 paper Friedman argued that economic models should not be evaluated by the realism of their assumptions, but rather by the correctness of their predictions.<sup>13</sup> Models are nothing but instruments for prediction. By making this claim, Friedman intended to liberate economists from concerns with unrealistic assumptions such as perfect information, perfect competition, expected utility maximization, etc. If models based on these assumptions make correct predictions about the phenomena they are intended to explain, then they are legitimate.

Clearly, the models ANs are based on often involve unrealistic assumptions. So why then do the founders of the approach not help themselves to Friedman’s methodological proposal? The main reason, I surmise, is Friedman’s identification of prediction with explanation. For Friedman, there is nothing more, that is apart from theory building and prediction, economists should be interested in. Throughout his paper he uses the term ‘explanation’ in quotation marks. For example: “Viewed as a body of substantive hypotheses, theory is to be judged by its predictive power for the class of phenomena which it is intended to ‘explain’”.<sup>14</sup> By explaining Friedman means nothing more than making predictions that turn out to be true.

But this is not the spirit of the historically-minded social scientists who advocate the AN methodology. For them explanation requires identification of causal mechanisms responsible for bringing about a phenomenon, not just its correct prediction. Moreover, the causal mechanisms must include facts about actors’ motives, beliefs, background and an explanation of how their collective action yielded the outcome in question. Indeed, Friedman has been widely criticized for relinquishing the traditional goals of causal explanation and the goal of building theories that correctly state the structure of reality.<sup>15</sup>

<sup>12</sup> Levi “Modeling Historical processes” 13; R. H. Bates, A. Grief, M. Levi, J. L. Rosenthal, B. R. Weingast, *Analytic Narratives* (Princeton NJ: Princeton University Press, 1998), 13 and fn 16.

<sup>13</sup> M. Friedman, “Methodology of Positive Economics” in D. M. Hausman (ed.) *The Philosophy of Economics* (Cambridge: Cambridge University Press, 1994) reprinted from M. Friedman *Essays in Positive Economics* (Chicago: University of Chicago Press, 1953).

<sup>14</sup> Friedman “Methodology of Positive Economics”, 184.

<sup>15</sup> H. Simon, “Testability and Approximation” in D. M. Hausman (ed.), *Philosophy of Economics*, (Cambridge: Cambridge University Press, 1994), 214–216.

But the puzzle remains: how can the AN proponents continue to advocate use of thoroughly idealized rational choice models while at the same time evaluating ANs on how the assumptions of their models “fit the facts”?

One possibility is to qualify this requirement by restricting the set of assumptions that must “fit the facts”. It is standard in the literature surrounding this question to draw a distinction between more and less important assumptions. For instance here’s historian and philosopher of economics Mary Morgan:

Passive irrelevant factors can and should be simplified away, and we can go with idealized accounts of motivations, provided we can give a rather more accurate representation of the behavior, situations or interactions relevant for our purposes. . . . The realism of assumptions does matter, but not of all of them at the same time.<sup>16</sup>

Unfortunately, Morgan does not tell us what makes relevant assumptions relevant and irrelevant ones irrelevant. But something like this distinction is at play in the debate between the proponents of ANs and Jon Elster.

### 3.3 *Elster’s Critique:*

In his *American Political Science Review* article he calls the AN proposal “a case of excessive ambition”.<sup>17</sup> His main criticisms of the approach are all an instance of this general issue: the models underpinning ANs often fail to adequately represent situations that they are supposed to represent. They fail to do so in several ways:

- they assume standard economic rationality (that is actors in the games obey expected utility maximization, different equilibrium conditions, and are only motivated by their “material interest”).<sup>18</sup> This flies in the face of widespread evidence from psychology and experimental economics that real people do not reason in this way. And it ignores the

---

<sup>16</sup> M. S. Morgan, “Model Experiments and Models in Experiments” in L. Magnani and N. J. Nersessian (eds.) *Model-Based Reasoning: Science, Technology, Values*, (Kluwer Academic/Plenum Publishers, New York, 2002), 56.

<sup>17</sup> J. Elster, “Rational Choice History: A Case of Excessive Ambition” *American Political Science Review*, 94/3, (2000) 685–695.

<sup>18</sup> *Ibid.*, 692.

very plausible possibility that people are swayed by emotions. This is not to say that emotions are incompatible with rationality, they may be and many researchers model emotions in rational terms.<sup>19</sup> But, however we choose to model emotions, “material interest” explanations can and do fail to. Moreover, Elster insists, we do not currently have a way of modeling failures of standard rationality in formal terms. “The social sciences today, . . . , cannot offer a formal model of the interaction between rational and non-rational concerns that would allow us to deduce specific implication for behavior”, Elster adds.<sup>20</sup>

- ANs model large collectives such as governments, clans, regional elites, etc as unitary actors with beliefs, desires and all the reasoning that individuals are assumed to do. At this point, Elster argues, “credulity breaks down”.<sup>21</sup> There is no evidence, nor any attempt to provide evidence, that collectives indeed behave as individuals.
- ANs make assumptions about actors’ intentions and beliefs without providing evidence that these intentions and beliefs really do obtain.
- ANs do not model uncertainty to the extent that it obtains in the phenomena they are supposed to represent. Models tend to assume full information when this clearly isn’t the case.

Each of these complaints is an instance of a larger claim: assumptions of models fail to “fit the facts”. Although Elster’s target are the ANs written by the original five authors (we shall shortly look at their response), most of his criticisms apply to Schultz’s AN as well. Schultz’s actors (the government, opposition and the target state) are assumed to act according to Bayesian Nash Equilibrium with all its complicated probability updating, and place a high premium on (re)election. They are all modeled as unitary actors, whereas in fact they are groups of many disparate individuals. Schultz does provide evidence that the intentions and beliefs of the actors are close to what the model assumes, so this criticism may not apply to his AN. Uncertainty-wise, although Schultz’s model is a model of incomplete information, it assumes that the government and the opposition party

---

<sup>19)</sup> For a recent example, see C. Bicchieri, *The Grammar of Society: The Nature and Dynamics of Social Norms* (Cambridge: Cambridge University Press, 2006).

<sup>20)</sup> Elster “The Case of Excessive Ambition”, 692.

<sup>21)</sup> *Ibid.*, 693.

have access to information about the costs of the war, that all actors know that each will play Bayesian Nash Equilibrium etc.

How important are these criticisms? Are the simplifications Elster is pointing to material to the models' explanatory ambitions? Or are they, in Morgan's words, merely passive and irrelevant factors. If so, how could we tell?

### 3.4 *De-idealization*

The traditional answer to this question is known as *de-idealization*. The proposal is to gauge the importance of particular unrealistic assumptions by replacing them with the more realistic ones and checking how it affects the prediction. If the predictions do not change substantively under the new set of more realistic assumptions, then the falsity of the old assumptions does not matter. So the assumptions that *can* be replaced with more realistic assumptions while preserving (to some degree) the predictions of the model relevant to explaining the target phenomenon do *not* need to be satisfied by the target phenomenon. This approach discussed by philosophers of science Ernan McMullin<sup>22</sup> and Daniel Hausman<sup>23</sup> (among many others) can be used to articulate conditions under which a model explains a phenomenon or some particular aspect of it:

(HM) A model explains a phenomenon if it satisfies all the assumptions of this model that cannot be de-idealized.

De-idealization certainly has place in the AN methodology – Schultz himself uses it to show that the assumption of office-seeking, for instance, can be relaxed to some extent without affecting the model's prediction. The problem is that de-idealization is often impossible. Not just often impossible. It's impossible in exactly the sort of situations that ANs are supposed to help us with. Rationality as expected utility maximization with all the strict assumptions it carries with it (precisely structured preferences, well-

<sup>22</sup> E. McMullin, "Galilean Idealization" *Studies in History and Philosophy of Science* 16/3 (1985), 247–273.

<sup>23</sup> D. M. Hausman, *The Inexact and Separate Science of Economics* (Cambridge: Cambridge University Press, 1992); D. M. Hausman, "Paul Samuelson as Dr. Frankenstein: When Idealizations Escape and Run Amuck" *Poznan Studies in the Philosophy of the Sciences and the Humanities, Idealization in Economics*. B. Hamminga and N. de Marchi, (eds). (1994) Amsterdam: Rodopi, 229–43.

behaved probabilities and godlike powers of computation) is a staple of ANs. But these are not assumptions that at the moment can be relaxed. Or in cases where they can be,<sup>24</sup> relaxing them does change the predictions of models substantively. (Indeed, the point of relaxing them is precisely to accommodate anomalous experimental results such as the endowment effect, the Allais paradox etc). In addition to these assumptions, it can also prove impossible to relax assumptions about various mathematical features of utility functions and probability distributions, such as their continuity and differentiability.

So, de-idealization cannot be the answer to Elster's criticisms, or at least not the full story. We need some other way of showing that the unrealistic assumptions the model makes do not matter. Without it HM fails as an account of model application.

### 3.5 *Response to Elster*

The advocates of the AN approach offer a different response to Elster's criticisms. Rather than attempting to show that the simplifications he is pointing out do not matter,<sup>25</sup> they insist that they are inevitable and must be accepted for pragmatic reasons. They invoke a methodological maxim that no theory should be rejected until we have a new theory. So even if rational choice modeling has its minuses, it is worth sticking to it because it is the only theory that is adequately developed at this moment. "Non-rational approaches to choice are not yet far enough along to provide an analytic approach that challenges or extends the traditional choice framework".<sup>26</sup>

It is certainly sound methodology to keep a theory that explains and predicts albeit imperfectly when no alternative is available. But is this the case in contemporary social science? Is it true that there are no decent explanations based on emotions, commitment and other non-rational factors?

---

<sup>24</sup> See work on behavioral economics and C. Starmer, "Developments in the Non-Expected Utility Theory: The Hunt for a Descriptive theory of choice under risk" *Journal of Economic Literature* 38, (2000), 332–382.

<sup>25</sup> There is some attempt to argue that the falsity of, for instance, the assumption of unitary actors does not matter (R. H. Bates, A. Grief, M. Levi, J. L. Rosenthal, B. R. Weingast "The Analytic Narrative Project" *American Political Science Review* 94/3 (2000) 696–702, 698). But the support for this claim is not that this assumption can be relaxed without serious consequences. Rather the argument seems to be that economists do it (i.e. treat firms as rational individuals), so political scientists should be allowed too.

<sup>26</sup> Bates et al. "The Analytic Narrative Project", 699.

It is pretty clear that there are. Some historians believe that Hitler's loss in the 1940 Battle of Britain can be attributed not to any rational decision making, but rather to a racial prejudice that worked in favor of Brits. Explanations of why advertising works routinely invokes emotional responses people have to various pricing and presentation tricks. Indeed, it is fair to say, that the whole discipline of psychology (and now behavioral economics) is dedicated precisely to this sort of explanations. These explanations are no more and no less mysterious than rational choice explanations.

Perhaps what the advocates of the AN approach mean is that these explanations are not sufficiently theoretically developed, that is, they lack the formal apparatus of rational choice theory that allows modeling of many different choice phenomena as utility maximizing. This is what they have in mind when they invoke an "analytic approach". That is certainly true, explanations based on emotions and biases tend to be less mathematized than rational choice explanations. But that does not make them worse off qua explanations. Why should the best explanations be those underwritten by a formal mathematized theory?

This is not to say that theories with sophisticated formal apparatus are no more desirable than more qualitative informal theories. Formal theories have an advantage when it comes to precision of predictions and assumptions, which helps to know what exactly the theory predicts and under what presuppositions. But this is just one advantage and it needs to be weighed against other criteria, most notably, empirical performance. And this advantage would only count if non-rational explanations suffered significantly from lack of precision of predictions and assumptions. If they don't and the explanation most supported by evidence is emotion-based rather than rationality-based, then no amount of formalisms in the rationality paradigm can make up for that fact. In many areas of social science, non-rational choice explanations are sufficiently well developed to compete with the rational choice ones.

So Elster's critique stands. As things look now, we have no evidence that the idealizations introduced by rational choice models do not matter. And there is no adequate justification for ignoring these idealizations in explanations.<sup>27</sup>

---

<sup>27</sup>) There may be good reason to ignore idealizations for other projects though. When our goal is to elaborate models as a part of theory development, we are plausibly engaged in an intellectual exercise rather than empirical work (Hausman *Inexact and Separate Science*), and in this case idealizations can be justifiably ignored.

### 3.6 *The Open Formulae Approach*

Does this mean that rational choice models, and idealized models more generally, should not be used in explanation? No, it doesn't. In what follows I propose how they should be used. Elsewhere I argue that this is indeed how they are used in the most empirically successful areas of economics.<sup>28</sup> However, the cost of this view is that it makes no sense to say that rational choice models explain phenomena to which they are applied or that they get confirmed in this process.

What often happens on the AN approach is the following: a model is presented which bears some, though not complete, similarity to the phenomenon in question. Although some of this model's assumptions are not satisfied it is treated as containing clues to the causal mechanism responsible for the phenomenon. However, since researchers may not have the evidence showing that the falsity of the unrealistic assumptions does not matter, they cannot treat the model as specifying this mechanism. Instead they treat the model as providing a framework for formulating a statement of a mechanism and they then look to the narrative to fill in this statement.

I call this view of models the Open Formulae view and defend it in a recent paper.<sup>29</sup>

On the Open Formulae view the model is not treated as making any claim, not even in the abstract. Rather the model functions as a template for a claim. It provides some of the categories which we can use to build a causal hypothesis. The open formula in our case takes the form of:

- (1) In a situation of type  $x$  with some characteristics that may include  $\{C_1 \dots C_n\}$ , a certain feature  $F$  causes a certain behavior  $B$ .

where  $x$  is a variable,  $F$  and  $B$  are property names, respectively, of putative causes and effects in the model, and  $\{C_1 \dots C_n\}$  are the conditions under which  $F$ s cause  $B$ s in the model. In an open formula,  $x$  is a free variable. That is, it is not yet specified, nor quantified over. An open formula needs to be filled in before it can make any claims. Once  $x$  is specified, we get a

<sup>28)</sup> A. Alexandrova, "Making Models Count" *Philosophy of Science*, 75(2008), 388–404 argues that this account makes sense of the now famous design of spectrum auctions based on a branch of game theory, auction theory.

<sup>29)</sup> Alexandrova, "Making Models Count".

causal hypothesis of the form “an F causes a B in a situation S”, where S is characterized by some conditions C. Without closing the open formula by specifying x, the open formula only gives us a *template* or a *schema*, or a *recipe* for a causal claim, rather than a fully fledged causal claim that is needed for explanation of a phenomenon. For example, one of the open formulae<sup>30</sup> of Schultz’s bargaining game is as follows:

- (2) In a situation of type x with some characteristics that may include {assumptions about information, motives, features of the good at stake, rationality of actors etc}, presence of the opposition party causes lower probability and greater credibility of government’s threats in an international crisis.

Note that the open formula allows that the conditions under which a result is derived in the model,  $\{C_1 \dots C_n\}$ , may or may not be the conditions under which Fs cause Bs in the real world. That is, the formulation leaves it open whether the hypothesis we construct on the basis of the open formula includes all or some of the model’s assumptions. For example, on the basis of the open formula (2), Schultz develops the following causal hypothesis:

- (3) Under conditions of political competition between office seeking government and an opposition party, open debate, and access to information about costs of a potential war (in the way these conditions obtained in Britain in 1898), a government’s threat in an international crisis such as the Fashoda crisis will be less probable and more credible.

Does it have to be part of (3) that the parties and the target state are fully rational, play Bayesian Nash Equilibrium, have well-behaved utility functions, draw their information from well-behaved probability distributions, etc.? Although all of these assumptions are required to derive the result of interest in the model, these assumptions may not form part of the causal hypothesis researchers formulate on the basis of the model. The model *constrains* but does not *determine* the hypothesis.

This freedom is essential when so many of the model’s assumptions are deeply unrealistic of the situation to which they are applied. Participants in

---

<sup>30</sup> I call the approach Open Formulae in plural because each model typically can yield several open formulae. I thank Don Garrett for making this clear to me.

the Fashoda crisis most probably did not satisfy most of the game's assumptions, and yet Schultz has made a good case that the opposition party's behavior did cause the threat to be more credible. So although the game suggests some of the categories which figure in (3), the mechanism which accounts for the outcome of the Fashoda crisis is different from the mechanism in the model. How could it not be? The reasoning that players employ in the game simply could not be realized by the real world players. This is why Schultz had to find a different set of conditions that realize the effect in question. These are whatever conditions that obtained in the Fashoda crisis that made it possible for the opposition party's behavior to cause greater credibility of the government's threat. (His ability to articulate such a fully fledged causal claim is what makes Schultz's AN a successful explanation of the Fashoda crisis.) Some of these conditions correspond to the assumptions of the game (plausibly, the office seeking of the parties and the presence of open political debate in Britain). But others do not.<sup>31</sup>

Note also how specific the claim (3) is. Although that is not always clear from the presentation of ANs, the hypothesis that these narratives confirm are not general hypotheses about the behavior of entities such as governments, opposition parties, etc. Rather the hypothesis follows closely the causal mechanism that the narrative identifies as underlying the explanandum. This is why the conditions under which democratic politics cause greater credibility of threats, as established by Schultz's AN, are very close to the conditions under which they caused it in the Fashoda case. How close? This is a difficult question to answer. In each case it is a complex empirical issue whether and how much the hypothesis established in one AN will generalize further. In my view it is best to err on the side of caution and formulate one's hypothesis very specifically. This is because the precise local conditions which obtained during the Fashoda crisis may very well turn out to be relevant to the operation of the postulated causal claim. Experimental economists who use game theoretical models to devise real

---

<sup>31</sup>) The distinction between the conditions under which the hypothesis is true in the real world and the assumptions needed to derive it in the model is particularly clear in experimental economics. When economists design institutions on the basis of their models they often cannot rely on the model's assumptions to tell them under what conditions a result that follows from a model will obtain in the real world (F. Guala, *Methodology of Experimental Economics* (Cambridge: Cambridge University Press, 2005), Alexandrova "Making Models Count").

world institutions learn that many more factors turn out to be relevant than thought at first. In case of auction design, for example, even small details of the software which implement the Federal Communications Commission spectrum auction are crucial for achieving the auction's economic objectives.<sup>32</sup> So when the auction technology tested in the lab needs to be transferred into the real world, researchers are very careful to reproduce the conditions of the lab as closely as possible. Of course a philosophical account cannot be expected to specify a priori just how specific should the hypothesis suggested by the model and established by the narrative be. This is a judgment call of every researcher.

To sum up, the methodology is as follows:

Step 1: Identify an open formula on the basis of a model by picking from the model's premises and conclusions the Fs and Bs of interest. These will correspond to the putative causes and the putative effects in the empirical phenomenon in question.

Step 2: Fill in x so as to arrive at a causal hypothesis of the form "Fs cause Bs under conditions C" where Fs and Bs match some aspects of the target situation.

Step 3: Confirm the causal hypothesis. This is done by finding a *material realization* of the causal relation in the hypothesis. A material realization is a material environment such that if it obtains, then an F causes a B.

To contrast this account with the traditional Hausman-McMullin account, here's an explicit statement of the role of models on my view:

(OF) A model applies to a phenomenon (for explanation) if there exists a material realization of a hypothesis based on the model's open formula in this phenomenon.

It is important to realize that on the OF view, the model merely inspires or informs the hypothesis that, if confirmed, does the explaining. The model cannot be said to do any explanation itself. Although it provides categories in terms of which the hypothesis about the real causal mechanism is for-

---

<sup>32</sup> C. R. Plott "Laboratory Experimental Testbeds: Application to the PCS Auction" *Journal of Economics and Management Strategy* 6/3(1997), 605–638.

mulated, it does not state this hypothesis itself. Remember that we still have no evidence that the simplifications used in the model do not matter, so we cannot claim that the mechanism proposed in the model is the mechanism that realizes the effect in the real world.

Although social scientists who practice the AN methodology largely proceed in accordance with the account I sketched here, they often treat the models as more than just the source of categories. Rather it is common to talk about models as if they describe the mechanisms, albeit in abstract terms, that can then be found instantiated in the real world. Idealized models are often credited with an ability to “capture the logic of issues”, “describe the heart of the problem”, “delineate the bare structure of social situations” to quote some unpublished sources overheard in seminars, talks and lectures. Both economists and political scientists who use economic methods assign models the ability “to capture the essence” of social phenomena<sup>33</sup> or to “show how people behave in various circumstances”.<sup>34</sup> Of course, it is also common to talk of rational choice theories as explaining real world phenomena, or being confirmed through tests. On my view, an open formula can be very fruitful in yielding causal hypotheses that explain and get confirmed, but it cannot itself explain and get confirmed. It is the causal hypothesis that does the explanation and if successful is confirmed. But the model is not sufficient to specify the hypothesis, it is merely a framework for it.<sup>35</sup>

Now we can return to our discussion of the criteria for evaluating ANs. Need the models assumptions “fit the facts” of the narrative? To some extent they must. Obviously if some assumptions are not satisfied, then the model cannot even be said to be *about* the phenomenon in question. For example, if Schultz’s model was applied to a polity that did not have open

---

<sup>33</sup>) Bates et al. *Analytic Narratives*, 12.

<sup>34</sup>) R. P. McAfee and J. McMillan “Analyzing the Airwaves Auction” *Journal of Economic Perspectives*, 10/1: (1996) 159–175.

<sup>35</sup>) So from the success of the hypothesis not much follows about the success of the model. Of course, it so happens that some models inspire more hypotheses that get confirmed than other models. For example, since John Harsanyi and Reinhard Selten revolutionized game theory by developing tools to model uncertainty, these models have yielded more hypotheses that got subsequently confirmed than models which did not incorporate uncertainty. This is evidence that these models are better at providing open formulae which serve as templates for statement of causal mechanisms. But exactly what this greater fruitfulness amounts to is a topic for another paper.

contestation of the governing office, such as an absolute monarchy, then the bargaining game would not even yield the right categories to construct a good causal hypothesis about this polity's behavior in an international crisis. However, once we know that there is such a minimal fit between the model and the situation in question, we do not need to insist that all its assumptions be satisfied by the situation, and we need not worry if we cannot de-idealize the model.

But there is a cost of this freedom. A casual similarity between the model and the situation does *not* give grounds to claim that the mechanism postulated in the model explains the target situation. Rather the model can be used to formulate a hypothesis about this situation and then this hypothesis needs to be confirmed. It can be confirmed by checking that the mechanism of the hypothesis is actually realized in the situation. Social scientists and historians do that by poring over primary sources in the archives, interviewing subjects, consulting statistical records, etc.

#### 4. Other Criteria of AN Evaluation:

Levi and her co-authors also propose three further criteria by which to judge ANs:

3. The model's prediction must fit the data.<sup>36</sup>
4. The explanation proposed by the AN must do better than alternatives.<sup>37</sup>
5. It is desirable that the explanation proposed be generalizable, rather than fit only the case under consideration.<sup>38</sup>

On the OF approach an AN provides a good explanation when it offers a causal hypothesis (informed by the model and the narrative) confirmed better than any other causal hypothesis. So the desiderata 3 and 4 are easily accommodated and endorsed by my account.

What about generalizability? Levi and her co-authors insist that the causal mechanisms identified in an AN needs to be "portable":

---

<sup>36</sup>) Bates et al. *Analytic Narratives* 17; Levi "Modeling Historical Processes" 14.

<sup>37</sup>) Ibid.

<sup>38</sup>) Levi "Modeling Historical Processes", 15; Bates et al. *Analytic Narratives*, 18.

By identifying the specific form of collective action problems, principal-agent issues, credible commitment, veto points, and the like, analytic narratives provide a way to suggest the characteristics of situations to which these apply and in what ways”.<sup>39</sup>

That is, an AN is better if it allows us to see the situation in question as an instance of a more general phenomenon. They also recognize that this requirement is in tension with the commitment of AN to account for “the particular puzzle in a particular place and time with a model and theory tailored to that situation”.<sup>40</sup>

The OF approach is neutral on this issue. On the face of it, generalizability of an AN does not necessarily make it a better causal explanation. If the historical puzzle under consideration does not fit under any of the standard rational choice categories, then the best explanation of this puzzle will not be generalizable in the way the proponents of ANs wish. But that does not make it any worse as an explanation.

Throughout this paper I have been assuming that a good explanation is one that identifies the causal mechanism responsible for the phenomenon in question – a view common among philosophers of social science.<sup>41</sup> On this view, generalizability is not a requirement on explanation, since a causal mechanism may be unique to a phenomenon. Of course, there is another philosophical account of explanation according to which generalizability is a crucial feature of explanation. This account is known as the unification theory of explanation.<sup>42</sup> According to this theory, good scientific explanations are unifications, that is, explanations that account for many different phenomena with a few argumentative patterns. For example, what was valuable about Newton’s mechanics as an explanation is that, relative to the Aristotelian framework, it unified celestial and earthly phenomena, thus, to use Philip Kitcher’s words, reducing “the number of facts we have to accept as ultimate”.<sup>43</sup>

---

<sup>39)</sup> Levi “Modeling Historical Processes” 15.

<sup>40)</sup> *Ibid.*, 14.

<sup>41)</sup> See fn 4.

<sup>42)</sup> P. Kitcher, “Explanatory Unification and the Causal Structure of the World” in P. Kitcher and W. Salmon (eds) *Scientific Explanation*, 410–505 (Minneapolis: University of Minnesota Press 1989).

<sup>43)</sup> *Ibid.*, 423.

Philosophers have discussed many different objections to the unificationist account of explanation.<sup>44</sup> But for our purposes only one point will suffice. The proponents of unificationist accounts of explanation claim that causal explanations in social sciences work in this way also – we seek to use a small number of variables to account for many different phenomena. ‘To account’ here means ‘to derive from an argument pattern’. But note that even accepting this view does not commit one to prefer a more generalizable game theoretical explanation to a less generalizable other kind of explanation. (I am assuming here that game theoretical explanations will tend to be more generalizable, than traditional historical narratives). Generalizability is not the only virtue of explanations even on the unificationist account. A generalizable explanation that does not imply the explanandum in an argument pattern, is no better, and perhaps worse, than a non-generalizable explanation that does imply the explanandum. It may well happen that a traditional narrative is more successful at that than a game theoretical model. So whether or not a more generalizable explanation is better than a less generalizable will depend on the details of each case, that is on which explanation correctly implies the phenomenon in question. Of course, a unificationist hopes that, as social science progresses, its explanations will become more and more generalizable.

If one does not accept a unificationist view of explanation, one might still value generalizability as a virtue of a research program, but not necessarily a virtue of explanation. Again, evaluation of research programs is a project separate from the question of this paper.

To take stock, I do not favor making generalizability a precondition of a successful AN. But if a generalizable statement of the causal mechanism an AN describes is available, this is a plus.

## 5. Need for Modesty?

The view defended here is that for ANs to provide successful explanations, the formal model at the heart of the narrative must inform the correct statement of the causal mechanism underlying the phenomenon in question. Contrary to what the original proponents of the ANs claim, for that it is not necessary that all the model’s assumptions be satisfied, nor that the

---

<sup>44</sup> For the latest, see Woodward *Making Things Happen*, 359–373.

causal mechanism be generalizable, nor that the model does the major bulk of explanatory work. It is, however, necessary that the model contains the right categories out of which to construct the statement of the mechanism, that is, that it correctly identifies some of the causes, effects and enabling conditions in question. I am in agreement with Levi and her co-authors that the model must be solved properly, that the events it identifies as causes and effects actually take place, that the statement of causal mechanism the model points to must survive competition with other explanations.

Where do these criteria leave us in terms of the viability of ANs as a distinctive approach to historical explanation? That is, are the criteria I identify as necessary for a successful AN so difficult to satisfy as to make the project impractical? Jon Elster appears to answer in the affirmative: the assumptions of rational choice models are so restrictive, that we should not expect AN to be successful in more than a handful of real world situations.<sup>45</sup> Hence he decries the ambition of the project as misplaced and unwarranted.

The OF approach allows idealized models to participate in explanation when its assumptions are not satisfied and when de-idealization of these assumptions is impossible. So long as it is not claimed that the mechanism proposed in the model is also the mechanism that explains the outcome in the world, the model can inform the explanation of a phenomenon in a meaningful way without mirroring this phenomenon. The approach is thus more permissive in what models we get to use and more restrictive in how much of the explanatory credit the model carries. So the OF view tempers Elster's pessimism about ANs, at the cost of relegating models to a lesser role. Thus understood, there is no reason to treat the AN approach as impractical.

However, even once our view of models is adjusted in the way I have argued it should be, modesty about the explanatory reach of the AN approach is still an appropriate attitude. Analytic narratives are only as good as the models which underpin them. If these models can only generate hypotheses which fit with the rational choice paradigm, then they will not be applicable elsewhere. There are puzzles in social science to which game theoretical tools have no interesting applications. As Theda Skocpol puts it:

---

<sup>45</sup>) Elster "The Case of Excessive Ambition".

Game-theoretic models are likely to make sense only of certain kinds of historical situations, where there really were sets of actors deliberately maneuvering in relation to one another. But not all historical processes of interest take this form. Often actors are blindsided by massive contextual changes or tweaked by unfolding processes they do not comprehend.<sup>46</sup>

In other words, a good historical explanation does not have to be, and sometimes should not be, an analytic narrative.

---

<sup>46</sup> T. Skocpol, “Commentary: Theory Tackles History” *Social Science History* 24:4, (2000), 669–676, 673–674.