The usefulness of game theory as a method for policy evaluation

Paper for the 10th EES Biennial Conference in Helsinki, October 3-5, 2012.

Leon M. Hermans*, Scott W. Cunningham, Jill H. Slinger

Delft University of Technology, Faculty of Technology, Policy and Management, Section Policy Analysis

P.O. Box 5015, 2600 GA Delft, The Netherlands

*Email: L.M.Hermans@tudelft.nl

Short abstract

Most of today's public policies are formulated and implemented in multi-actor networks. Logically, this has consequences for evaluation methods and approaches. Game theory has long been around as a method that supports a rigorous analysis of the interaction processes among actors. However, so far, it has not been widely applied in the evaluation field. Hence, questions regarding the usefulness of game theory as an evaluation method remain pertinent. This paper explores the usefulness of game theory as a method for evaluations in a networked society, based on an evaluation of coastal policy implementation in the Netherlands.

Key words: Game theory, multi-actor systems, coastal management, evaluation methods

Introduction

In a networked society, policy formulation and implementation is shaped by the interaction between multiple actors. This multi-actor characteristic is not new, neither for implementation (e.g. Bardach, 1977) nor for policy formulation (e.g. Rittel and Webber, 1973). However, the visibility and importance of these networked characteristics are growing. How does the evaluation profession respond to this rise of the networked society and the multi-actor characteristics of modern governance? One can address this question by assessing existing evaluation methods and concepts for their usefulness in evaluating network governance interventions (Hertting and Vedung, 2012). One can also look at the development and adoption of *new* methods in evaluation of network governance interventions. In this case, two types of responses are visible: a process response, and an analytical response.

The emergence of the process response led to the development of a constructivist evaluation approach as described by Guba and Lincoln in 1989. This process response resulted in the increased incorporation of participation, consultation, interaction and similar activities in evaluation. An analytical response, however, has been less well-developed. In part, the process response formulated by Guba and Lincoln was accompanied by a particular constructivist paradigm and methodology. Within this paradigm, there is a growing toolkit that offers analytical support for facilitating participatory processes of joint sense-making (e.g. Moret-Hartman et al, 2011; Hermans et al, 2012). However, besides analytical support for participatory constructivist processes, also analytical tools and methods are available that help to understand processes among multiple actors.

There are a good many of such methods that combine analytical rigour with practical feasibility, as explored for situations of ex-ante policy analysis (see Hermans and Thissen 2009 for an overview). Examples in evaluation include applications of social network analysis to evaluate platforms and advocacy groups (Drew et al., 2011), stakeholder analysis to support evaluation design (Bryson et al, 2011), and game theory (Langbein, 1994; Niklasson, 1996). Despite these examples, attention for this type of analytical support appears to lag behind other areas of methodology development in evaluation.

Therefore, the focus here is on one specific analytical approach: game theory. Game theory is of interest as it offers “a theory of behavior in strategic situations” (Greif, 2006: 18). It is connected to the groundbreaking work of Von Neumann and Morgenstern and John Nash in the 1940s and early 1950s. Game theory thus has long been around and it is commonly used in economics and political science. Yet, in evaluation literature few applications are reported to date. This paper explores if game theory would be a useful part of the analytical toolbox in this regard. This issue is addressed through a short literature review and a case report of the application of game theory to evaluate implementation of coastal policy in the Netherlands.

The potential usefulness of game theory for evaluation

Game theory explains the outcomes of interactions based on rational choices by individual participants. Game theory formalizes a game metaphor to study the strategic interactions among actors, where actors have to coordinate their behaviour with that of others, and where actors anticipate on the behaviour of others to decide on their own strategies. Basic concepts in game theory include actors (players), who each have a number of possible strategies or actions to follow, whereby the strategies chosen by each actor determine the outcome of the game. Payoffs represent the value of these outcomes to different actors (Straffin, 1993). Games are governed by rules.

Game theory models can simulate a range of real-world games quite accurately and game theory holds promise as an analytical framework for policy analysis and institutional analysis (e.g. Ostrom, 1990; Scharpf, 1997; Greif, 2006). However, game theory also has important limitations. In the literature, most attention has gone to the assumptions that underlie game theory: a limited number of fully rational actors, involved in an isolated game, with complete information on the game and the rules of the game, as well as their own preferences and utility functions. This has led to some specific branches of game theory that further develop one or more of these limiting assumptions.

Another limitation is that the practical applications of game theory – and similar models that assume rational actors – are hindered by various manifestation of bounded rationality and irrational behavior. Games may be assumed to exist as a theoretical construct, but observing these games, and measuring various parameters in practice, is difficult. Clearly, assessing pay-off structures for various actors is difficult – a problem known also in the larger field of economics for instance related to the importance of stated versus revealed preferences. Strategic games involve strategic behavior. Participants are not likely to be open about this, and may not be willing to share information with evaluators. Also, given that social actors are capable of adapting and learning, capturing social processes fully in formal (game theoretic) models is nearly impossible.

These limitations are important to note, but do not necessarily mean that game theory could not be a useful addition to the toolbox of evaluators. For one, policy and planning debates have a frontstage and a backstage. They are debated in public based on rational merits – even if behind the curtains personality clashes influence the outcome of decision-making (Hermans & Thissen, 2009: 815). And, as all approaches that assume rationality in decision-making, making explicit what is known and applying logical structures to make use of the information that *is* available, is widely accepted as better than simply groping in the dark. The popularity of logic-models in project evaluations is a point in case. Finally, as game theorists like Rapoport (1970) have argued in the past: the worth of game theory is not so much in predicting what will happen (or, for that matter, explaining the necessity of past events), but rather, in limiting the space of possible outcomes. Often, this is already worthwhile.

In evaluations there are few applications of game theory. At least, a literature search on game theory in three leading evaluation journals (Evaluation; Evaluation & Program Planning; Evaluation Review) yields only a handful of papers. Of those, some articles merely mention game theory, but do not specifically address or report about the usefulness of game theoretic methods for evaluations. Two articles are more interesting here. Langbein (1994) uses game theory in an evaluation of the enforcement of regulatory programs. Niklasson (1996) reports an application of game theory to evaluate the implementation of university reform policies in Sweden. These articles by Langbein and Niklasson both conclude that game theory methods bring useful benefits. However, they have not been followed up since their publication in the mid-1990s, meriting a renewed exploration of game theory and its usefulness for evaluation.

Case: implementing coastline management in the Netherlands

This section contains a concise description of a case, the details of which can be found in a working paper by Hermans et al. (2011). Rather than including numerous references to case materials and sources here, the interested reader is referred to this more elaborate working paper for underlying sources.

The Dutch are internationally renowned for their efforts to keep dry feet and protect their low-lying lands against flooding. The Delta works are well known, but less visible internationally are the soft sandy infrastructures that protect the Dutch coasts. The majority of the Dutch coastline consists of sandy beaches and dunes. For centuries, this coastline has been subject to gradual erosion and inland movement. Since the early 1990s, however, an official coastal policy came into working that was to halt coastal erosion and maintain the Dutch coastline at its 1990-reference position. Implementation of this policy was done through annual sand nourishment programs. Every year, six to twelve million m^3 of sand were to be distributed along the sandy coast, at places where this was considered necessary in order to maintain the coastline.

The distribution of this sand was to be decided through an annual procedure in which national government (by means of Rijkswaterstaat, the Dutch department for Public Works and Water Management), provincial government, municipalities, water boards and nature organizations were represented. They discussed annual sand nourishment programs in so-called “Provincial Consultative Bodies”. Consultation of these regional bodies, which were chaired by the provinces, was required by law. This procedure was designed by experts from Rijkswaterstaat, partly to ensure sufficient support for the annual sand nourishment programs, and partly to improve the fit of these annual programs with regional needs and priorities.

Formalizing actors and interest in games in the Provincial Consultative Bodies

One can summarize the discussions and decisions in the Provincial Consultative Bodies as a game. In this game, five actors can be distinguished, as summarized in Table 1. The play of this game was regulated through official procedures. Rijkswaterstaat was a first mover in these games, as it proposed an annual nourishment program. This proposal was based on technical calculation rules, filled with annual monitoring data about the actual position of the coastline.

Table 1. Actors and their interests in the Provincial Consultative Bodies

	Actor	**Interest**
M	Municipality	Local economy, tourism, recreation
N	Nature	Nature conservation and ecological rehabilitation
R	Rijkswaterstaat	National coast safety, “delta safety”
W	Water board	Local management: safety
P	Province	Process: acceptable outcome for coastal zone

In game theory, players are assumed to act rationally, meaning that they are assumed to take those actions that are likely to maximize the expected utility of the outcomes. Therefore, a game theoretic analysis requires an estimation of the utility of different outcomes for different players. For this utility modeling, three lines of evidence were used. There was engineering evidence of the utility of the players based on physical system monitoring data and models. Interviews with key players were used for the face validation of the game. Expert opinion was used as basis for value elicitation.

The cooperative game theory approach attempts to explore the fundaments of collective action. The approach uses as an atomic units of analysis the creation of value by individuals, groups and coalitions. Consequently, the elicitation of value requires many hypotheticals – what would happen if these two coalitions could align their interests?

An important part of this explication of utility and value involves action on the part of the other coalitions. In this work we adopted the maximin aggregation method. This means to review for each coalition: how much value can this coalition generate assuming a hostile response from the other parties?   Value, in this particular game, was assessed by reasoning primarily from safety interests. This means that the game was analyzed from a particular angle. This was warranted by the observation that all players in the game acknowledged safety as the first and foremost concern in coastline maintenance; Basic safety should not be traded-off against recreational or nature interests.

The value potential of all coalitions, measured in delivered safety, was rated on a scale from 0 to 100. Utility measurement lacks an absolute scaling, so we may arbitrarily set the value to scale between zero and one hundred. Note that this exercise does not guarantee that the grand coalition of all players delivers the most value; this is a matter for judgment and investigation on a case-by-case basis. Indeed, many serious policy issues surround coalition building processes where it is difficult or impossible to bring a coalition on-board without a global loss of value. Such situations are known as sub-additive games.

Game in Characteristic Function Form

The game as evaluated is set out below. The operator ν{} should be read as "the value of coalition X is." As previously noted we select four players, or idealized interest groups, to capture the key features of this game. The players are represented by the letters R, W, M and N. P (province) was omitted, as this actor officially acted as a neutral chair. In reality, provinces sometimes were perceived to align with municipalities and their interest in the local economy and recreation.

The game as estimated by the research is given below. Note that the point of this particular part of the analysis is not to get the exact numbers of each coalitions value right. Rather, quantification here serves to give an indication of relative values of coalitions, based on inputs from experts and informed by interviews.

$$\nu\{\phi\} = 0$$

$$\nu\{R\} = 94 \qquad \nu\{W\} = 8 \qquad \nu\{M\} = 13 \qquad \nu\{N\} = 19$$

$$\nu\{RW\} = 100 \qquad \nu\{RM\} = 63 \qquad \nu\{RN\} = 69 \qquad \nu\{RM\} = 38 \qquad \nu\{WN\} = 44 \qquad \nu\{MN\} = 6$$

$$\nu\{WMN\} = 31 \qquad \nu\{RMN\} = 69 \qquad \nu\{RWN\} = 75 \qquad \nu\{RWM\} = 75$$

$$\nu\{RWMN\} = 81$$

Solution Concepts: Core and Shapley value

The goal of cooperative game theory is to provide strategic advice concerning the play of a game. In cooperative games, we hope to describe who should cooperate with who, and how much these parties should sacrifice in support of the common interest. However, arriving at such "solutions" for cooperative games is not straightforward. There are multiple means for cooperation and there are different ways to assess whether cooperative outcomes are fair, sufficient or efficient. We adopted two particular solution concepts in our analysis. First, we used a modest proposition for examining design solutions, one which is known as the "core." The core considers only the space of solutions which are in the individual and group interest. (See Rapoport 1970 for a discussion of these and other solution concepts for cooperative games).

Analyzing the design space using the solution concept of the core, reveals that there is no core to the game; There is no solution which meets the needs of individual and group rationality. All players agree on the importance of safety as the main interest. They do not agree on the additional values such as those related to recreation/tourism, nature and dune preservation. Basically, the additional non-safety values put on the game by the players {M,N,W} cost additional budget. However, the available budgets are not sufficient to cater to these values. In practical terms, this was illustrated by the frequent demands for additional (small) sand nourishments by municipalities, and the standard response by Rijkswaterstaat: additional nourishments may be incorporated, but not on Rijkswaterstaat's budget.

Another solution concept that was explored for this game was the use of the Shapley value. This is a mediation solution which attempts to reward players according to what they bring to the table. However, also with this solution concept there will be dissatisfaction. For instance, the value of the game for the coalition of Water board, Municipality, and Nature v{WMN) = 31 (out of 100). If now Rijkswaterstaat would join them, the value of the resulting grand coalition would be v{RWMN} = 81. This means that Rijkswaterstaat would bring 50 units of value to the table. If it was to join an existing coalition of the other three players, it could thus expect a maximum share of 50 units. However, on its own, Rijkswaterstaat would be able to secure v{R}=94 units of satisfaction on its safety interests. The assumption here being that, on their own, they would not have to cater to any of the additional demands. Thus, Rijkswaterstaat has no incentive whatsoever to join any coalition other than the coalition with Water boards. The safety value of this coalition is estimated at v{R,W}=100. What water boards bring to the table, is their specific knowledge of local and regional safety concerns, allowing for fine-tuning of sand nourishments for safety.

Implications of games repeated over time

It was not possible within the modeled game to meet all the additional non-safety values. The safety values were accepted by all players as the main values. The rules of the game gave further influence to Rijkswaterstaat, as both the first mover in the game, and the final decision-maker. Only within these boundaries, was there some room for individual players to have some influence.

These sand distribution games were repeated every year, for several years. As a result, some of the other players may have grown dissatisfied. Players indicated in interviews that they had an increasing feeling that the discussions were outdated rituals rather than meaningful joint decision-making. Therefore, it is understandable that these players, notably municipalities, nature organizations and provinces, gradually moved the policy debate to other venues, where they expected better opportunities to realize outcomes that would be favorable for them.

This is visible in the diminishing role in the more recent years of the Provincial Consultative Bodies for Coast. As of 2009, with a new Water Bill, the provincial consultative bodies no longer have an official role in the programming of nourishments – they need to be informed, but their consent is no longer required. Annual sand nourishment programming is still being discussed at this venue, but the big decisions on coastal management are taken elsewhere.

For instance, the most visible decision in Dutch coastal management in recent years has been the construction of a so-called "sand motor" in 2011. This sand-motor was a pilot project whereby an artificial appendix to the coast was created, basically a large volume of sand, which should gradually erode over a period of twenty years or more, to nourish beaches and dunes further north along the coastline. The actors responsible for safety (the water board and Rijkswaterstaat), stressed that the effects of the sand motor were uncertain, and that, despite the large volume of sand, 20 million m^3, still additional sand nourishments might be necessary if the sand would not be released at the right time and locations to maintain the base coastline. In the decision about the sand motor, it is clear that Rijkswaterstaat, which once designed the sand distribution game in a way that allowed for national control, was not in control anymore. Considerations other than flood safety and efficiency in providing safety were more influential to explain the outcome of decision-making.

Discussion of the usefulness of game theory in this case

Game theory has helped to see that the implementation of an earlier coastal policy decision resulted in a game that had no core, meaning that not all individual demands could be satisfied. Rijkswaterstaat was in control, while other players could not be certain about the outcomes of interest to them. This helps explain why these other players eventually lost interest and looked elsewhere for possibilities to influence coastal management. But is this useful? Does it add something to existing evaluation reports in this case?

To a certain extent, it does. As any theoretical lens, game theory illuminates phenomena that would otherwise remain less clear. In this regard, it is useful to understand how the original design of the implementation procedure logically explains the observed outcome. Rather than merely signaling disappointment, such understanding enables the design of new procedures and "games" that enable more effective negotiations and stipulate new forms for collaboration that offer more possibilities to align mutual interests. Because even if other players have moved key coastal management debates to other venues, they still remain interested in safety. Very much so even – and they do realize they need the expertise and resources of Rijkswaterstaat as part of the effort to secure safety. In the end, sidelining their expertise is not in their interest either.

Past evaluations signaled that there was some dissatisfaction with annual procedures. Decisions regarding non-safety issues were considered intransparent. However, they did not signal that part of this was due to a lack of interest in non-safety issues on behalf of the dominant player in the game, and that part of it was due to a lack of information, with which to make decisions more transparent. Safety, equated with the prevention of coastline erosion, was well understood, due to past monitoring efforts of Rijkswaterstaat backed by a community of coastal scientists. The value of stable, broad or narrow beaches for recreation and the local economy was less clear. Also the impacts of sand nourishment on nature were ill-understood. This focus on the role of knowledge and information in games does not a priori follow from the application of a game theory lens, but game theory does help to understand that different players have different needs for information to operate effectively in games. Hence, limiting the information base and knowledge development efforts to those issues of interest to one of the players, is, again, not addressing the concerns of other players in the game.

Evaluations that were done earlier by and large used the same basic information as a starting point, often also combined with interviews with key stakeholders. Yet, this game theoretic evaluation results in another story. As such, game theory enables one to build stories that could be true – and from which lessons can be drawn for future decisions. For this, value is in the structure provided by game theory, identifying the core elements and relations that explain interactions as games. The (semi) quantification associated with much of game theory helps sharpen and test the resulting storyline, but to us does not seem to be its main feature. Also in our case, the numbers with coalition values remained crude and imprecise.

Conclusion

Game theory offers a lens that enables one to analyze policy processes in which multiple actors are involved. As such, it can inform evaluators and evaluation users about these processes and their design. Game theory is not unique though. Other tools may offer lenses that are equally effective – notable the field labeled here as actor analysis contains many more of such lenses.

Lenses are influential. If you look for games, you will see games. If you look for networks, you will see networks. If you look for mechanisms, you will see mechanisms – and if you look for impact, you will see impact (or sometimes, the lack thereof). Therefore, this paper is not a call for the application of game theory everywhere. However, it is a call for a conscious choice of analytical lenses and theoretical concepts. When interactions among multiple actors are an imported part of the policy or program that is to be evaluated, neglecting lenses that help explain such interactions would be a bad choice.

Acknowledgements

We would like to acknowledge financial support from the Next Generation Infrastructure Foundation under the project "Games and Learning in Delta Infrastructures". Also, we would like thank our collaborators in this project, especially Jan Mulder and Ankie Bruens from Deltares and Marja Menke from Arcadis, as well as interviewees and participants in various workshops.

References

Bardach, E. (1977). *The Implementation Game: What Happens After a Bill Becomes a Law*. The MIT Press, Cambridge, Massachusetts (second printing, 1979).

Bryson, J. M., Patton, M. Q., & Bowman, R. A. (2011). Working with evaluation stakeholders: A rationale, step-wise approach and toolkit. *Evaluation and Program Planning* 34(1): 1-12.

Drew, R., P. Aggleton, H. Chalmers, K. Wood (2011). Using social network analysis to evaluate a complex policy network . *Evaluation* 17(4): 383-394.

Greif, A. (2006). *Institutions and the Path to the Modern Economy: Lessons from Medieval Trade*. Cambridge University Press.

Guba, E. G., & Lincoln, Y. S. (1989). *Fourth generation evaluation*. Newbury Park: SAGE Publications

Hermans, L.M. and W.A.H. Thissen (2009). Actor analysis methods and their use for public policy analysts. *European Journal of Operational Research* 196: 808–818.

Hermans, L.M., Cunningham, S.W. et al. (2011). *Games and Learning in Delta Infrastructures: Coastline Management in the Netherlands*. Working Paper. Delft University of Technology / Next Generation Infrastructures, Delft, the Netherlands.

Hermans, L.M., A.C. Naber, B. Enserink (2012). An approach to design long-term monitoring and evaluation frameworks in multi-actor systems - a case in water management. *Evaluation and Program Planning* 35 (4): 427 – 438.

Hertting, N. and E. Vedung (2012). Purposes and criteria in network governance evaluation: How far does standard evaluation vocabulary takes us? *Evaluation* 18(1): 27-46.

Langbein, L.I. (1994). Estimating the impact of regulatory program enforcement. Practical Implications of Positive Political Theory. *Evaluation Review* 18(5): 543 – 573.
Niklasson, L. (1996). A Cultural Revolution in the Universities. The Possible Uses of Rational Choice Models. *Evaluation* 4(3): 278 – 293.
Ostrom, E. (2005). *Understanding Institutional Diversity*. Princeton University Press, Princeton, NJ.
Rapoport, A. (1970). *N-Person Game Theory: Concepts and Applications*. University of Michigan Press: Ann Arbor.
Scharpf, F. W. (1997). *Games Real Actors Play: Actor-Centered Institutionalism in Policy Research*, Westview Press, Boulder, CO.
Straffin, P.D. (1993). *Game theory and strategy*. The Mathematical Association of America. Washington, DC.

Author biographies

The three authors are faculty members at the Faculty of Technology, Policy and Management of Delft University of Technology.

Leon Hermans has a PhD in policy analysis from Delft University of Technology. His current research focuses on methods and approaches for monitoring and evaluation to support learning in multi-actor systems. Key areas of interest are water and environment, as well as international development.

Scott Cunningham has a D.Phil. from the Science Policy Research Unit of the University of Sussex. He applies methods of decision analysis, content analysis and game theory to study questions of technology management.

Jill Slinger has a PhD in Applied Mathematics from the University of KwaZulu Natal. Her current activities focus on the role of (model-based) knowledge in supporting interactions between people, science and policy in the governance of water and coastal systems.