Introduction

In accordance with what it does as a regional organisation, the European Union (EU) has a ‘big external footprint’ (Dinan, 2010: 483). Next to having nearly 140 delegations in individual countries, the EU is represented in many international organisations (Nugent, 2010: 393). Relations with these multilateral institutions, including the World Trade Organization (WTO), the World Bank, and the North Atlantic Treaty Organization, are well established. Nonetheless, the added value of EU interventions in multilateral institutions is not self-evident. The EU is the first to take into account the potential shortcomings that go with claiming a bigger role in international organisations. One of the objectives of the Lisbon Treaty is to address deficiencies associated with the EU’s external relations. In terms of greater policy coherence, effectiveness and visibility of the Common Security and Defence Policy (CSDP) for example, much is expected of the Lisbon Treaty (Directorate General External Policies of the Union, 2008). Especially the establishment of a High Representative of the Union for Foreign Affairs and Security Policy and the European External Action Service (EEAS) are thought of as important new steps for the EU’s external policy. Being double-hatted as the chair of the Foreign Affairs Council and the vice president of the European Commission (EC), the High Representative – assisted by the EEAS – is tasked to bring more coherence in the EU’s external relations.

With the increasing engagement of the EU in international politics and the concomitant maturing of foreign policy institutions the growing relevance of assessing EU performance in the foreign policy area is beyond doubt. Properly executed evaluation research on EU interventions in multilateral institutions is all the more important if one takes into consideration that, although the EU regards effective multilateralism to be a key policy objective, this investment has not always been translated into growing influence in the institutions at hand and that there is substantial variation in EU performance in different organisations (Jørgensen, 2009). A key question to be answered is: what does performance mean? How should one assess the performance of the EU in a multilateral environment? In this chapter the claim is made that simply taking over the dominant method of measuring performance by comparing the EU’s policy objectives (inputs) with policy effects (outcomes) is unsatisfactory for measuring EU performance in multilateral organisations. An alternative approach is proposed in which EU performance in international organisations
takes place on the level of diplomatic interventions rather than on the level of the EU as a whole, and in which performance is assessed in the context of the successive stages of the evolving multilateral policy process. Doing so, the too often ignored complex causality of the multilateral political environment in which EU diplomats operate is fully taken into account.

The chapter starts with a section on performance measurement of international organisations in general. It is shown that many evaluation models use as a starting point linear causality to measure the performance of international organisations by linking objectives directly to effects. The second section proposes an alternative evaluation model based on complex causality in which the black box of the negotiations is opened. The third section elaborates on what conditions are needed to assess EU diplomacy in a multilateral environment. Three key-features of multilateral diplomacy are identified which illustrate the complex causality of multilateral environments. In sections four and five, a model is presented that was developed in earlier research to evaluate national diplomatic interventions in multilateral decision-making arenas (Kleistra and Willigen, 2010). The model is adapted to measure the performance of the EU diplomatic interventions in multilateral institutions. It consists of a framework for analysing diplomatic interventions in multilateral settings and context-bound evaluation criteria. Finally, in the last section some suggestions for the way forward for thinking of EU performance in multilateral institutions are made.

Linear outcome performance

Performance as a concept refers both to the completion of a certain task and the way in which this task was achieved. In that sense, a distinction can be made between outcome performance (reaching objectives) and process performance (the ability to reach objectives) (Gutner and Thompson, 2010). Measuring the process performance of an international organisation means doing a process evaluation, or an evaluation of the internal dynamics of implementing organisations, their policy instruments, their service delivery mechanisms, their management practices and the linkages among them (DAC/OECD, 2002). Measuring outcome performance comes down to doing a policy evaluation, or measuring performance as the completion of a certain task. The latter is the focus of this chapter.

Outcome performance measurement has a longstanding tradition in the area of policy evaluation research as such. Since the late 1950s, policy analysts have generated a variety of (chiefly quantitative) methods to study the functioning of national policy programmes and their effects. These include among others survey-analysis, cost-benefit analysis, cost-utility analysis, cost-feasibility analysis, quasi-experimental designs, interrupted time-series design, non-equivalent comparison group design and multiple regression analysis (Worthen et al., 1997). Outcome performance measurement is often based on the assumption that there is a linear relationship between specifically defined objectives of an organisation and the policy effects. As a result, models and measurement methods aiming to compare policy objectives directly with policy effects dominate the theory and practice of policy evaluation (Howlett et al., 2009; Khakee, 2003).

Linear outcome performance measurement starts with identifying the policy objectives (input), continues with analysing the policy (output), and ends with studying the effects of the policies (outcome) while comparing these with the objectives. An illustrative example is the approach taken by Lusthaus. Lusthaus distinguishes four elements of policy performance: effectiveness (the achievement of objectives), efficiency (an appropriate cost structure), relevance (able to adapt to changing circumstances and thus remain relevant), and financial viability (the inflow of financial sources is larger than the outflow) (Lusthaus et al., 1999). Understood in this way, outcome performance is synonymous to goal achievement or effectiveness. The emphasis on linear outcome performance measurement can for a large part be explained by the close linkages between the
development of evaluation policy and public sector reform (Wollman, 2003). In particular, the New Public Management movement of the 1980s and 1990s, with its emphasis on effectiveness and results, led to an evaluation practice which compared policy goals directly with policy effects. The key assumptions and beliefs of linear outcome performance evaluations are aptly summarised by John (1998: 22): ‘Through distinguishing between policy goals and output/outcomes policy analysts are able to find out if policy intentions turn into reality, and when policies are successes or failures.’

The insights that result from this evaluation practice are meant to support decision making and public accountability. This does not only go for national governments, but also for international organisations. The United Nations (UN) for example, expects evaluation to serve ‘as an input to provide decision-makers with knowledge and evidence about performance and good practices’ (United Nations Evaluation Group, 2005). The EU explicitly links evaluation to the decision-making cycle through ex ante, interim evaluation, ex post evaluations. Different from (liberal-democratic) national settings, the EC (as the EU’s principal evaluating body) states that evaluations serve mainly the decision-making process of the Commission, rather than accountability towards EU citizens (Stern, 2009). According to the EC, the purpose of evaluation is: ‘to provide policy-makers with input about the impact and effectiveness of activities planned and carried out’ (Commission of the European Communities, 2006: 3).

In the area of development politics, the focus on the results of the interventions of national states or international organisations finds its base in an almost universally accepted linear outcome performance measurement model. In 1991 the Development Assistance Committee of the Organisation for Economic Co-operation and Development (DAC/OECD) has developed for DAC-members a set of evaluation principles for evaluation and results-based management. These principles were refined into five criteria that since have been widely used in the evaluation of development initiatives. The criteria have a strong resemblance with the above mentioned Lusthaus approach. The criteria are: efficiency (measures how economically inputs are converted to results), effectiveness (looks at the extent to which objectives were achieved), impact (looks at the wider effects of the policy), sustainability (likelihood of continuation of the benefits produced by the policy after the policy initiatives have been completed) and relevance (the extent to which the objectives are consistent with the priorities of the target group, recipient and donor) (DAC/OECD, 2002). The EC, which is an observer to the OECD, works alongside the 34 member states in the preparation of texts and participate in discussions on the OECD’s work programme and strategies (OECD). It fully embraced the DAC/OECD evaluation guidelines in its evaluation practice. This has meant that to date the EU defines evaluation narrowly as linear outcome performance. This is clearly reflected in how the EC defines (policy) evaluation, namely as a ‘judgement of interventions according to their results, impacts and the needs they aim to satisfy’ (Commission of the European Communities, 2000: 2).

Not only the EU itself, but also outside observers of EU performance often measure performance in terms of linear outcomes. They use the EU’s declared objectives and investigate the extent to which these objectives have been met. A recent example from the wide selection from which we now can draw is the Evaluation of EC-funded Mine Actions 2002–7 which was executed in 2010 by the Geneva International Centre for Humanitarian Demining (GICHD, 2010). The mine action evaluation reflects the broad application of the DAC/OECD-evaluation framework, since the mine action programmes represent EC-initiatives on the interface of security, humanitarian assistance and development cooperation. What is more important for our argument is that the mine action evaluation at the same time shows, that by focusing on linear outcome performance, all the initiatives that are less tangible, but constitute an even valuable contribution to the issue at hand, are not included in the final assessment. The focus of the evaluators is exclusively on the
financial input of EC as a major donor in multilateral mine action programmes, while the input of EC-delegates in international landmine negotiation arenas – multilateral diplomacy as a second foreign policy tool – is completely ignored.

**Complex outcome performance**

Today, the mainstream literature on public management and policy analysis is full of criticism of linear outcome performance as such (John, 1998; Howlett et al., 2009; McConnell, 2010). Listing again the general shortcomings of the linear outcome performance approach adds little to this knowledge. In fact, the criticism reflects a long-standing gap in the field of evaluation theory between positivist models and approaches and what Guba and Lincoln (1989) have mentioned the fourth generation of policy evaluation. Fourth generation evaluation theorists reject the dominant rational evaluation approach, both its assumptions and its methods. Their scientific point of departure is not positivism but social constructivism. With it, a shift is made from causal explanation to the logic of policy deliberation. Problems, objectives and solutions are no longer perceived as given facts, but as social constructs (Fischer, 1995).

The gap represents two philosophical extremes on a continuum. At the social-constructivist extreme is the interpretative policy analyst who attempts to ‘understand’ or to ‘make sense’ of social phenomena by trying to get inside the actor’s realm of subjective experiences and to reconstruct the actor’s own action motives, while relying mainly on qualitative methods such as public panel series, participatory observation, focus group design and discourse analysis. At the other extreme is the professional policy evaluator who generally adopts the positivist approach. These evaluators at least feel uncomfortable by making any use of the social-constructivist methodological toolbox and the alternate role for the evaluator that comes with it (Weiss, 1998). The model presented in this chapter is within the positivist tradition. What is relevant to the future of EU evaluation is the understanding that adopting a model which aims at assessing the performance of the EU in international organisations according to a linear causal logic is counterproductive. In the case of relying on a linear outcome performance model at least two serious difficulties arise.

First of all, a linear outcome performance measurement works best when policy objectives are unambiguously clear, do not change over time, and are set in a static policy environment. However, policy objectives are often vague, they do change during the policy implementation, and they might become irrelevant or counterproductive due to a changed policy environment (Jørgensen, 1998: 89–90). This applies even more when it comes to pursuing policy objectives in multilateral settings. In the foreign policy literature, it is therefore emphasised that foreign policy, or in this case, the diplomatic interventions in international multilateral settings, is too complex to be framed in a linear input-output-outcome model (Brighi and Hill, 2008). What goes for foreign policy in general also goes for the actions of the EU in multilateral organisations: it is not only to be expected, but also desirable that EU’s policy objectives are formulated in general and abstract terms. The interests of European citizens would not be served if the EU embarked on establishing very specific but infeasible foreign policy goals, rather than adapting appropriately to the international negotiation dynamics. It is precisely this feature of policymaking under these circumstances that makes it difficult to assess whether the EU objectives have been met.

Second, linear outcome performance assumes that it is possible to distinguish linear causal relationships between the policy inputs of an actor and the policy outcome. However, in multilateral settings it is hardly possible to distinguish the contribution to a specific outcome of one actor from another by way of a linear outcome performance analysis. Assessment of EU foreign policy on the basis of linearity often ignores the complex causal dynamics of these multi-actor settings. As the EU’s contribution is implemented through multilateral arrangements, the
quality of its foreign policy performance should for the larger part be judged by the ‘standards of collective effect and collective benefit’ (Webber and Smith, 2002: 101). The outcome of foreign policy is not merely determined by the input of a single actor, but also by the environment in which the actor operates. In other words there is ‘a dialectic interplay between the actor’s own strategy on the one hand, and context on the other hand’ (Brighi and Hill, 2008: 119). As such, European success in external relations can hardly be conceived of as a positive achievement of *ex ante* formulated European policy objectives. Instead, the implementation of foreign policy through multilateral diplomatic channels should be seen as ‘a continuous act of negotiation on several fronts, with no final resolution of the central issues’ (Webber and Smith, 2002: 101). The insight that foreign policy implementation seldom leads to a final resolution of complex issues is aptly illustrated by the classic definition of diplomacy as ‘*négociation continue*’. At its core, diplomacy is about permanent ‘communication between officials designed to promote foreign policy’ (Berridge, 2010: 1).

In order to avoid these difficulties, outcome performance in multilateral institutions should not be understood as a linear causal process in which objectives can directly be related to outcomes. Instead, outcome performance is characterised by complex causality in which the outcomes flow from ‘the convergence of several conditions, independent variables, or causal chains’ (George and Bennett, 2005: 212). Consequently every policy input should be assessed within its own proper context (Karns and Mingst, 1992). This means that, the concept of outcome performance needs to be unpacked if we want to learn about best (and worst) practices. In that way, conflicting policy goals might become clear as well as differences in performance by one international organisation across several issue areas. Unpacking can be done by looking at specific diplomatic interventions dealing with a specific policy issue. In the next three sections, a complex outcome performance model is presented for evaluating the performance of EU diplomatic interventions in multilateral institutions. First, the key features of multilateral diplomacy are elaborated upon, because they determine the complex environment in which diplomatic interventions take place. Secondly, a heuristic framework is presented that enables us to analyse diplomatic interventions through process tracing. And thirdly, policy evaluation criteria are formulated that take into account the context within the interventions take place.

**Key features of multilateral diplomacy**

Diplomacy, as an instrument of foreign policy, is defined as: ‘the conduct of international relations by negotiation rather than by force, propaganda, or recourse to law, and by other peaceful means (such as gathering information or engendering goodwill) which are either directly or indirectly designed to promote negotiation’ (Berridge, 1995: 1). As the definition suggests, negotiation is the key activity of diplomats. That being said, negotiations may be supported by the threat or use of force in which case we speak of *coercive* diplomacy. Since on the whole, coercion is not used in multilateral diplomacy, for the purpose of this chapter we ignore this element and focus on diplomacy as non-coercive negotiations. EU negotiations in multilateral institutions are characterised by constraints on the participants’ individual ability to control outcomes. Outcomes of multilateral negotiations are not in the first instance determined by the negotiation skills of individual countries, but by ‘a management process with actors seeking to reach agreements through a process of adjustment’ (White, 2005: 401).

Three key features of multilateral diplomacy can be identified. The first feature is that it is time-absorbing. It may take many years before decision-making processes result in actual outcomes. In the course of time policy objectives might change. This is because results are achieved by means of small steps that do not radically deviate from the status quo. Most multilateral arenas are characterised by routines, norms, habits and procedures that are difficult to change and often
hamper swift and successful negotiation. Therefore, negotiators have to adapt their goals and aspirations to the specific features and proceedings of multilateral negotiations. From this it follows that any complex outcome performance model will have to deal with the aspect of policy change that follows naturally from the pace and progress of decision-making at the international level.

The second feature of multilateral diplomacy is that the power of an actor to achieve a specific policy objective in a given multilateral decision-making arena is limited by definition. Each actor has its own interests and objectives, as well as different opinions on the urgency of problems and valuable solutions to these problems. This may result in some actors preferring the status quo and obstructing any attempt to formulate, decide on and implement new policy initiatives, while other actors may support change in certain directions. This means that policymakers who participate in these settings are constantly put under pressure to formulate ambitious policy objectives, while in realising these objectives they are dependent on other actors. So, any complex outcome performance measurement model will have to incorporate the ability of other actors to block policy initiatives.

The third feature is that the process of interaction of multiple (equal) actors in multilateral decision-making arenas implies that the effectiveness of the performance of any actor, whether it be a national state or the joint actions of 27 states under the umbrella of the EU, cannot be judged by simply linking (parts of) the outcomes of international negotiations to national or regional policy ambitions. These outcomes are first and foremost the result of concerted action and have to be judged as such. Assessment of the effectiveness of any actor in multilateral settings is further complicated by the very fact that participating in a given multilateral decision-making arena – and thus opting for concerted action – entails that its effort in this arena is meant not only to benefit the actor’s interests, but also the interests of the international community at large. This means that every policy initiative has to be assessed bearing in mind the context in which the initiative was accomplished, the policy strategy and the extent to which the result is still acceptable as compared with the policy objective.

The features listed above apply to multilateral diplomatic negotiations in general. It can be argued that multilateral negotiations carried out by the EU are even more complex than multilateral negotiations carried out by a single state. The EU’s relations with international institutions are defined by a high heterogeneity of preferences. As a multilateral system itself, EU policies result by definition from bargaining (Alesina et al., 2005). Bargaining not only takes place between the 27 member states, but also between the different EU institutions. Although the Lisbon Treaty is meant to make the EU’s external relations more coherent and effective, the EU will remain a heterogeneous actor. This is clearly shown by the co-existence of the Permanent Presidency and the High Representative/Vice-President of the Commission (both having a role in representing the EU externally) and by the membership of the EEAS which will be a mixture of permanent staff of the Council Secretariat, the Commission, and seconded diplomats from the European capitals (Vanhoonacker and Reslow, 2010; Dinan, 2010).

This persistent heterogeneity should be taken into account when doing performance measurement of the EU’s external relations. At the same time, it should not stop there. In the end the outcome performance of the EU is determined by what the EU actually does in the multilateral institutions. That being said, given the nature of multilateral negotiations and the complex environment in which they take place, the effectiveness of the EU’s multilateral diplomacy cannot validly be assessed by simply linking policy goals and outcomes. How then, could the relations of the EU with international organisations be evaluated on the level of diplomatic interventions? First, a framework for analysis is needed that takes into account complex outcome performance as an ongoing process. Secondly, context-bound evaluation criteria are formulated. This model will be presented in the sections that follow.
Framework for analysing diplomatic interventions in multilateral institutions

Taking into account the features of policy change, the power to block initiatives, and the emphasis on collective effects and benefits in multilateral negotiations, the so-called barrier model is a suitable heuristic tool for analysing diplomatic interventions. Summarising, the model is based on the assumption that actors aim to achieve specific objectives in a given decision-making setting. When doing so, they meet resistance from other actors involved. Those who are able to overcome the obstructers or convince others win the obstacle race. An actor who decides to strive for a specific goal needs to be prepared and able to overcome a large set of barriers (impediments) throughout the decision-making process – not only during the activity of putting an item on the agenda, but also when developing alternative solutions, gaining support for these solutions and monitoring the actual implementation of decisions already taken. Hence, any actor committed to achieving certain objectives participates in an obstacle race by definition: 1) it has to turn wishes into demands; 2) it has to turn demands into issues; 3) it has to influence the decision-making process; and 4) it has to influence the implementation process in its favour.

As a heuristic tool the barrier model is basically a stages model which allows process tracing. Stages models are widely used in public policy analysis with their added value being the ability to ‘capture’ complex decision-making processes and therefore their usefulness as a heuristic device (Hill, 2005: 20–21; John, 1998: 36). Like stages models in general, the barrier model assumes that, in a policy process, a number of fairly distinct stages follow each other in a fixed order. It is precisely this feature that makes the model appropriate to be used for an evaluation of EU diplomatic interventions in multilateral organisations. Negotiations in multilateral negotiation settings operate according to fixed patterns and rigid procedures. Diplomats concerned with multilateral interventions often regard these frameworks as indispensable, but also – and not infrequently – as an obstacle to achieving results. The various phases or stages are determined by the general course of events regarding the issue at question. The clearly defined stages of the multilateral decision making process does not only allow us to map in an appropriate way the otherwise elusive, or maybe random-like set of actions of the EU or any other individual participant. It allows us also to assess the successive EU actions in its proper context.

Taking the barrier model as a framework for analysis, diplomatic interventions in various phases of decision-making for different issues under discussion can be structured into five sections: general course of events, objectives, results, interventions and impediments. The five sections of the framework for analysis of the evaluation are defined in the box below.

Framework for analysis of EU diplomatic interventions in multilateral arenas

- **General course of events** describes the decision-making process regarding the issue in question on an international level. Substantial steps taken in the multilateral arena relating to that particular issue are listed, together with a brief review of the current state of affairs.
- **Objectives** present an overview of the EU objectives regarding the issue in question. This will offer insight on how objectives have changed (or not) in the course of time.
- **Results** give an overview of which EU objectives have been achieved and which have not.
- **Interventions** focus on the kind and frequency of interventions undertaken by the EU delegation.
- **Impediments** describe the different barriers in the different policy phases encountered by the EU negotiators.
This framework enables a process tracing effort in which the diplomatic interventions are causally linked to the outcome of the negotiation process (George and Bennett, 2005). Thus, like with any stages model, linear causality is preserved, because the model adopts the idea that policies follow a certain sequence. At the same time, it differs from the linear outcome performance measurement described above, because the black box between the policy objectives/input and the policy effects/outcome is opened. Doing so, so called policy feedbacks or loops might be detected in which events influence decisions (objectives) taken at an earlier stage (John, 1998: 23; Brighi and Hill, 2008: 120) Thus, using this framework one can determine which causal links there are between specific policy objectives of the EU and the outcome of the negotiations, while taking into account the specific interventions and the impediments which influenced the results. Having done that, an evaluation should be made which is based on the notion that multilateral negotiations are determined by complex causality rather than linear causality.

Evaluation criteria

It is thus our claim that an evaluation of the EU’s performance in international organisations should take into account the contextual features of multilateral diplomacy. What criteria are we talking about? First, the intervention should support the objectives formulated at the European decision-making level. The evaluation criterion connectedness refers to the degree to which the interventions are in line with the policy objectives as formulated by the EU as a whole or its specific foreign policy institutions and the degree to which the interventions echo the level of ambition expressed in the policy objectives. Following this logic, an intervention may be connected, partly connected or not connected to the EU’s policy objectives.

Second, the intervention should be conducive to, or stimulate the coming about of consensus of opinion in the international decision-making community. The evaluation criterion responsiveness looks at the degree to which interventions are conducive to the larger international decision-making process. An intervention may be in line with the already prevailing views or ideas of the international community on the matter. An intervention may aim at overcoming the differences between two extreme positions, or it may aim at adopting a specific position. Such a position may also be very innovative or likely to be controversial, but nevertheless effective by offering a breakthrough in a deadlock. Following this line of thought, an intervention may be responsive, partially responsive or not responsive to the international decision-making process.

Third, the intervention should be executed at a convenient moment. Most interventions are not usable in every phase of the decision-making process. When does one undertake what intervention? The evaluation criterion timeliness is concerned with assessing the usefulness and effects of different interventions in different phases of the decision-making process. Depending on the circumstances, an intervention may be too early, timely or too late. The lengthy process of coming to common positions is one of the pitfalls of EU decision making structure. Since timeliness requires experienced and smooth-working foreign policy institutions it is to be expected, that, despite the implementation of the Lisbon Treaty, the execution of interventions at convenient moments is likely to remain the biggest challenge for the EU.

Fourth, in order to be influential the intervention should have a reasonable scope. The scope of interventions may differ in three dimensions. The first and second dimensions are the action radius and the directness of the used instrument, respectively. When an intervention is direct and when it has a large action radius, the scope is considered to be greater than when the intervention is indirect and has a small action radius. For example, supporting, sponsoring or initiating a
resolution evidently has a wider reach than a démarche and thus has a large action radius. At the same time, a resolution is a highly indirect intervention in comparison to a démarche. The third dimension is the level of involvement of the actor executing the intervention. An actor proactively involved by chairing an international meeting or an international commission is likely to be more influential in the international decision-making process than an actor supporting a discussion paper or merely participating in a meeting. As such, a continuum of involvement can be created, which ranges from passive involvement to proactive involvement. The relationship between the three dimensions of scope and the most widely used intervention instruments is summarised in Table 7.1. All three dimensions together form the scope of the particular intervention. Scope is qualified as being small, moderate or large.

After having scored the EU diplomatic interventions on the criteria of connectedness, responsiveness, timeliness and/or scope, an assessment must be made whether the low scores are justified given the context the intervention took place. For example: a policy issue on which the EU claims to fulfil a proactive role, but which at the same time is characterised by interventions of a small scope in a particular phase of the decision-making process may be justified by the fact that another issue was deemed to be more important at the time. It may have been necessary to opt for one issue rather than another, because of the scarcity of time and resources. The four criteria which have been discussed above are summarised in Table 7.2.

### Table 7.1 The three dimensions of scope related to intervention instruments

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Action radius</th>
<th>Directness</th>
<th>Involvement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participation in meetings</td>
<td>Large</td>
<td>Indirect</td>
<td>Passive</td>
</tr>
<tr>
<td>Undertaking a démarche</td>
<td>Small</td>
<td>Direct</td>
<td>Active</td>
</tr>
<tr>
<td>Support proposal/resolution/paper</td>
<td>Large</td>
<td>Indirect</td>
<td>Active</td>
</tr>
<tr>
<td>Sponsor proposal/resolution/paper</td>
<td>Large</td>
<td>Indirect</td>
<td>Active</td>
</tr>
<tr>
<td>Initiate proposal/resolution/paper</td>
<td>Large</td>
<td>Indirect</td>
<td>Proactive</td>
</tr>
<tr>
<td>Chairing meetings</td>
<td>Large</td>
<td>Direct</td>
<td>Proactive</td>
</tr>
<tr>
<td>Organising meetings</td>
<td>Small</td>
<td>Direct</td>
<td>Proactive</td>
</tr>
</tbody>
</table>

### Table 7.2 Evaluation criteria

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Definition</th>
<th>Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Connectedness</td>
<td>Criterion refers to the degree to which interventions are in line with policy objectives as formulated in European foreign policy and the degree to which interventions follow ambition expressed in policy objectives.</td>
<td>Connected</td>
</tr>
<tr>
<td>Responsiveness</td>
<td>Criterion looks at the degree to which interventions are conducive to the international decision-making process.</td>
<td>Responsive</td>
</tr>
<tr>
<td>Timeliness</td>
<td>Criterion looks at the degree to which the intervention took place at an appropriate moment.</td>
<td>Too early</td>
</tr>
<tr>
<td>Scope</td>
<td>Criterion refers to the action radius and directness of the intervention in combination with the level of involvement of the actor carrying out the intervention.</td>
<td>Large</td>
</tr>
</tbody>
</table>
Evaluating EU performance: the way forward

When assessing the performance of the EU in multilateral organisations it does not seem undue that getting consensus within the EU is already a big negotiation achievement by itself. Some even argue that this can be seen as the key success of European foreign policy (Dover, 2009: 256). However, with the growing importance of the EU as an actor on the international stage it will be inevitable to go beyond the mere creation of European consensus as an indicator of performance and focus on the added value of EU actions when measuring outcome performance. Within multilateral organisations such as the WTO and the UN, the EU is increasingly acting as an autonomous actor next to its member states (Mortensen, 2009; Laatikainen and Degrand-Guillaud, 2010). Consequently, in the near future it will become more important to assess to what extent the EU is successful in pursuing its policies in multilateral institutions.

It is our claim that linear outcome performance measurement, by comparing the EU’s policy objectives directly with policy effects, is unsatisfactory. The nature of the new EU external policy institutions and the very nature of the work they increasingly will be confronted with, requires us to go beyond the linear outcome performance practice. An alternative approach of complex outcome performance was presented in which the complexity of multilateral negotiations is taken into account. Complex outcome performance ideally avoids quick and dirty conclusions about EU (in)effectiveness. It would typically avoid accusations of ineffectiveness because outcomes do not match with the initial objectives. It would also avoid claiming that the EU is ineffective only because there is a lack of consensus within the EU on a certain policy issue. Instead, it would recognise that a compromised policy goal – however weak – could lead to a successful EU intervention. Instead of focusing on a lack of consensus with the EU as the explanation for a policy failure (as is often done), performance can better be explained by looking at the negotiation process between the EU and other actors. The EU might have a limited policy objective because of a lack of consensus on a certain issue, but be very successful during the negotiations to get this limited objective reached.

Notes

1 Parts of this chapter are based on earlier published work (Kleistra and Willigen, 2010). The authors thank the participants in the conference ‘Europe in Global Governance: Performance, Reform, Power’ at Rapallo in the autumn of 2010 for their useful remarks on a draft version of this chapter.

2 See for a more elaborate explanation of the barrier model, Kleistra and Willigen, 2010.

References