

15. A good example is the final report on the first evaluation exercise on mutual legal assistance in criminal matters, OJ C 216 of 01.08.2001.
16. Article 22(5) of the Framework Decision 2006/783/JHA, OJ L 328 of 24.11.2006.
17. This is one of the most substantial 'blocks' of JHA legislation so far, comprising essentially Council Directive 2005/85/EC of 1 December 2005 on minimum standards on procedures in Member States for granting and withdrawing refugee status, Council Directive 2004/83/EC of 29 April 2004 on minimum standards for the qualification and status of third country nationals or stateless persons as refugees, Council Regulation (EC) No 343/2003 of 18 February 2003 establishing the criteria and mechanisms for determining the Member State responsible for examining an asylum application, Council Directive 2003/9/EC of 27 January 2003 on minimum standards for the reception of asylum seekers, and Council Directive 2001/55/EC of 20 July 2001 on minimum standards for giving temporary protection in the event of a mass influx of displaced persons.
18. In 2005, for instance, regarding practices of cooperation in the fight against trafficking in human beings, drug-trafficking, and the counterfeiting of the Euro (Eurojust 2006: 41–6).
19. New Article 70 TFEU (Treaty on the Functioning of the EU) provides for 'the objective and impartial evaluation of the implementation of the Union policies referred to in this Title by Member States' authorities', and new Article 68 TFEU provides for the European Council 'to define the strategic guidelines for legislative and operational planning within the area of freedom, security and justice'.

11

The Role of Evaluation in Experimentalist Governance: Learning by Monitoring in the Establishment of the Area of Freedom, Security, and Justice*

Olivier De Schutter

1 Introduction

The decentralized approach adopted in democratic experimentalism encourages the subunits of a federal system to devise their own solutions to the regulatory problems they face. But local experiments will benefit the other subunits only if they are evaluated, according to scales which are at once flexible enough to accommodate the novelty of experiments that work, and sufficiently robust in order to provide the adequate incentives to the subunits. Through evaluation, the subunits should be encouraged to take part in a collective search for solutions which can be replicated elsewhere; and they should be discouraged from experimenting in ways which create negative externalities, which could lead to calls for the imposition of standards from above.

This chapter seeks to contribute to our understanding of the role of evaluation mechanisms in the architecture of democratic experimentalism, by focusing on the rise of evaluation in the establishment of the Area of Freedom, Security, and Justice (AFSJ). The establishment of an AFSJ between the Member States of the European Union is based on the idea that national courts and administrations, as well as law enforcement authorities, should cooperate with one another, in particular by exchanging information and by mutually recognizing judicial decisions in civil and criminal matters. Such cooperation presupposes that the Member States share a set of common values, which

* This paper was prepared as part of the REFGOV project (no. CIT-CT-2005-513450) funded by the 6th EC Framework Programme on Research and Development coordinated by the CPDR-UCL.

include the fundamental rights recognized in EU law.¹ It may also demand, in certain cases, the approximation of the national legislations of the EU Member States, as may be required for such cooperation.²

The standard explanatory framework for the progressive establishment of an AFSJ between the EU Member States is thus based on two complementary propositions. First, mutual recognition, the 'cornerstone' of establishment of the European criminal area, or other forms of cooperation between States in the field of law enforcement, presuppose a high degree of compliance with fundamental rights. The national authorities of the Member States should therefore refuse to cooperate where this would risk violation of these values: this is not only prescribed by the EU Treaty; it is also reiterated in the instruments implementing the principle of mutual recognition in a variety of domains since the concept was introduced at the Tampere European Council of 1999.³ Second, where there remain obstacles to mutual recognition or other forms of inter-state cooperation, due to diverging standards of protection of fundamental rights and to the resulting lack of 'mutual trust' between national authorities, harmonization may be required. Such approximation of national legislation might serve, in particular, to raise the overall level of protection of fundamental rights, and create the 'mutual trust' between the Member States which cannot merely be presupposed. We thus seem to witness in this area what neo-functionalists would see as confirming their view of the logic of European integration: 'positive integration', in the form of harmonization of national legislation, accompanies 'negative integration', especially where the latter takes the form of mutual recognition of national rules or decisions; harmonization is the result of spillover from the abolition of barriers to cooperation between the Member States.

This chapter challenges that classic narrative. It does not question the appeal of the standard view to the institutional actors. Nor does it underestimate the weight of the analogy to the establishment of the internal market in the mental representations of these actors—and, to that extent at least, the validity of the neo-functional logic of a complementarity at work between negative and positive integration. Rather, the aim of this chapter is to draw attention to a competing logic, which may help us move beyond the debate between mutual recognition and harmonization. This competing logic is a combination of evaluation and collective learning: by setting up evaluation mechanisms and by mutually observing one another the EU Member States not only can create the mutual trust on which their cooperation depends, but they also can make progress together towards identifying the precise content of this new field of European integration.

This chapter documents the emergence in the EU of a logic of evaluation and learning. This logic is an alternative to the balancing between positive integration and negative integration, since it is reducible to neither and opens up a third avenue through which to achieve coordination. It is also a complement to those

classic tools of integration, since we will use the latter better once we equip ourselves with the search devices which evaluation mechanisms can constitute. The approach is procedural, rather than substantive. It acknowledges that we cannot know, in advance of developing cooperation between EU Member States, which obstacles such cooperation may face, and which measures should be adopted in order to remove these obstacles. But the alternative is not necessarily to rely on a purely *ad hoc* construction of the AFSJ, guided by the priorities of national political agendas and the rhythm of crises occasionally drawing the attention of policymakers to certain, previously unidentified or underestimated, problems. Rather, evaluation mechanisms described here should be conceived as search mechanisms, which should allow us to identify, on a systematic basis, what steps are required to achieve progress towards the establishment of the AFSJ. Through the tool of evaluation, the ends of the AFSJ are constantly redefined as a result of developing the means to achieve it, and its shape is being discovered at the same time as the area itself is being invented.

The chapter proceeds as follows. Section 2 briefly describes the standard view of how progress should be made towards the establishment of an AFSJ, by analogy with the establishment of the internal market. That standard view opposes mutual recognition (or some other form of cooperation built on mutual trust) to harmonization, understood as the adoption of common standards. A trade-off between diversity and unity is implicit in this standard view: the risks of divergent approaches by the Member States, insofar as they could threaten mutual recognition, are to be countered by the imposition of uniformity—harmonization from above. In part, the logic of evaluation presented here is useful because it may allow us to escape such a trade-off. In order to describe how this alternative logic has emerged, Section 3 describes how the Member States have gradually come to realize that they needed to mutually evaluate themselves in a variety of fields, such as external border control, combating terrorism, or the administration of justice. The review of evaluation mechanisms in these areas illustrates that they fulfil at least five distinct objectives. They may serve to monitor the implementation of EU laws and policies by the Member States; ensure a feedback on those laws and policies themselves; contribute to collective learning, on the basis of local experiments; enhance mutual trust between the Member States; and finally, stimulate democratic deliberation both at national and at European level. Section 4 examines these different functions of evaluation, and asks whether they can be reconciled in a single model. Section 5 concludes.

The logic of evaluation presented here may be seen as subverting the neo-functional logic of which it is a potential competitor. But it may also be seen as a necessary complement to that neo-functional logic itself. Indeed, the establishment of an AFSJ between the EU Member States cannot avoid constantly questioning the content of such an area itself, and in particular, the relationship between mutual recognition and harmonization

(or, more broadly, between mutual cooperation and the definition of common standards) in its progressive establishment between EU Member States. By establishing such an area, the Union seeks to achieve a balance between conflicting goals—free movement of persons on the one hand, a high level of security on the other hand—and it does so by means which are as much competitive as they are complementary: mutual recognition and cooperation to the fullest extent possible, accompanied by approximation of national legislation or the development of common standards where necessary. Where the balance is to be struck, which degree of legal approximation should accompany mutual recognition and according to which sequence this should happen are left for us to discover as we move towards the fulfilment of the objective set by the Treaty. Evaluation mechanisms therefore may be conceived as search devices which will enable us to better understand this objective, in the very process of implementing it. This chapter explores the potential of this logic of evaluation, including both monitoring and learning.

2 The standard view

The standard view sees the establishment of an AFSJ in the European Union as the search for an adequate equilibrium point between negative and positive integration. This view identifies a deep structure in the process of European integration, analogizing the establishment of the AFSJ in the 2000s, to the establishment of the internal market in the 1980s. Mutual recognition is the rule, based on the premise that the national authorities of all Member States can be trusted to comply with the same set of publicly agreed upon values. Harmonization is the exception, but it constitutes the preferred remedy where mutual confidence breaks down, whether or not for objectively justifiable reasons: it constitutes the other horn of the dilemma. In this view, fundamental rights fulfil, in the establishment of the AFSJ, the same function as the 'mandatory requirements' famously put forward by the European Court of Justice in the *Cassis de Dijon* judgment of 1979 where, for the first time, the concept of mutual recognition was introduced in the law of the internal market.⁴ In *Cassis de Dijon*, the Court took the view that products should be allowed to be sold into any other Member State provided that they have been lawfully produced and marketed in one of the Member States, but it acknowledged that 'obstacles to movement within the Community resulting from disparities between the national laws relating to the marketing of the products in question must be accepted in so far as those provisions may be recognized as being necessary in order to satisfy mandatory requirements relating in particular to the effectiveness of fiscal supervision, the protection of public health, the fairness of commercial transactions and the defence of the consumer' (par. 8). One generation later, the European Court of Justice was asked

in the landmark case of *Gözutok and Brügge*⁵ whether the national courts of the Member States should be obliged, under the *non bis in idem* (double jeopardy) principle enshrined in the Schengen Agreement,⁶ to recognize that further prosecution is barred after the accused has arrived at a settlement with the prosecuting authorities of another Member State than the one where he or she is facing criminal charges. The Court explicitly noted that mutual recognition was not conditional upon the harmonization of criminal procedures across the Member States. Instead, said the Court, the 'necessary implication' of the *non bis in idem* principle is that 'the Member States have mutual trust in their criminal justice systems and that each of them recognises the criminal law in force in the other Member States even when the outcome would be different if its own national law were applied' (par. 33).⁷ The analogy between the concept of mutual recognition in the internal market and its function in AFSJ is clear: the latter thus emerges as a 'market of fundamental rights', in the words of Advocate General Ruiz-Jarabo Colomer.

Just as the AFSJ inherited the concept of mutual recognition from the law of the internal market, it has remained hostage to a strangely binary form of thinking characteristic of the 'new approach' to market integration. This standard view builds on the idea of a grand alternative between mutual recognition (in its many incarnations) and harmonization (see also Peers 2004). Only a few years ago, the approach of the Commission still offered a clear illustration of this. 'Mutual recognition', according to the Commission, 'is a principle that is widely understood as being based on the thought that while another state may not deal with a certain matter in the same or even a similar way as one's own state, the results will be such that they are accepted as equivalent to decisions by one's own state' (European Commission 2000a: par. 3.1). However, mutual recognition thus understood 'rests on mutual trust and confidence between the Member States' legal systems', which may have to be 'enhanced' by certain harmonization measures: 'Differences in the way human rights are translated into practice in national procedural rules [...] run the risk of hindering mutual trust and confidence which is the basis of mutual recognition' (European Commission 2003d: title I.7). The Commission stated thus, in 2005, that '[t]he first endeavours to apply the [mutual recognition] principle, in particular with the European arrest warrant, revealed a series of difficulties which could to some extent be resolved if the Union were to adopt harmonisation legislation [aimed at] ensuring that mutually recognised judgments meet high standards in terms of securing personal rights' (European Commission 2005d: par. 3.1).⁸

Nor is this view applicable only to mutual recognition in the criminal justice field. A similar dialectic is for instance currently at play as regards the protection of personal data processed by law enforcement authorities, in the establishment of the area of freedom, security, and justice. Just as the harmonization of the protection of personal data in the internal market was seen as a condition of mutual recognition of the relevant national legislations,⁹ the development

of common rules on the protection of personal data is considered a condition for the exchange of information between the law enforcement agencies of the Member States under what came to be called the principle of availability. This principle means that, 'throughout the Union, a law enforcement officer in one Member State who needs information in order to perform his duties can obtain this from another Member State and that the law enforcement agency in the other Member State which holds this information will make it available for the stated purpose, taking into account the requirement of ongoing investigations in that state' (European Council 2004: par. 2.1). In other terms, information available in one Member State should be made available to the authorities of any other Member State, just as if these were authorities of the same State: 'The mere fact that information crosses borders should no longer be relevant' (European Commission 2005e: par. 2.2). This 'principle of availability' is currently codified in the proposal for a Framework Decision on the exchange of information (European Commission 2005f).

The principle of availability plays in this field the role which the principle of mutual recognition plays in the field of judicial cooperation in criminal matters. It presupposes the mutual trust which should exist between the Member States' national authorities. But it is also a technique through which leverage may be exercised in favour of the adoption of common standards in order to strengthen mutual trust (de Biolley 2006: 194). Indeed, its implementation requires that all Member States ensure a high level of protection of personal data, thus justifying the high level of trust which this principle presupposes between the national authorities of the different Member States (European Council 2004: par. 2.1). However, the 1995 Data Protection Directive does not apply to the processing of personal data effected in the course of state activities in areas of criminal law or matters falling under Title VI EU.¹⁰ Moreover, while all the EU Member States are parties to the Council of Europe Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data of 1981, the principles set forth in this instrument are expressed at a relatively high level of generality, and certainly does not ensure the same level of protection as, for instance, the 1995 Data Protection Directive. Therefore, almost simultaneously to proposing an instrument implementing the principle of availability, the Commission put forward a proposal for a Framework Decision on the protection of personal data processed in the framework of police and judicial cooperation in criminal matters (European Commission 2005g). This was encouraged by the European Parliament (2005) and welcomed by the European Data Protection Supervisor (2006). It illustrates perfectly the standard narrative, in which the need to adopt 'flanking measures', aimed at improving the level of protection of fundamental rights in the EU Member States, appears as a logical—and unavoidable—counterpart to the lowering the barriers to mutual recognition or exchange of information.

There is, of course, an inherent tension underlying the dialectic described in the standard view. Insofar as it justifies mutual recognition, mutual trust is presupposed by the very fact that each Member State has agreed to consider decisions adopted by the authorities of any other Member State as equivalent to decisions adopted by its own authorities: such a presupposition was, for instance, central to the reasoning of the *Güziotok* and *Brügge* judgment of the European Court of Justice. But insofar as it justifies, instead, the approximation of national legislations, the concept of mutual trust appears rather as a precondition for establishing the area of freedom, security and justice on the principle of mutual recognition (De Schutter 2005; Weyembergh 2004: 339). In this second perspective, mutual trust is not to be taken for granted: it has to be created. The desire to strengthen mutual trust may therefore justify the approximation or the harmonization of legislations as a measure accompanying mutual recognition.

Whether this latter function of the notion of mutual trust corresponds to the original understanding of mutual recognition may be doubted. The EU Treaty provides in Article 31c, that common action on judicial cooperation in criminal matters shall include, *inter alia*, 'ensuring compatibility in rules applicable in the Member States, as may be necessary to improve such cooperation'. But this remains a particularly vague formulation. And when the concept of mutual recognition was originally put forward, with the evident purpose of achieving in the establishment of the criminal area what had been achieved in the internal market during the late 1980s, this was presented not as a lever to promote the further approximation of national legislations in this area, but quite to the contrary as a *substitute* for harmonization. The United Kingdom Presidency document of 1998 which initially presented the idea of mutual recognition stated: '... a possible approach, comparable to that used to unblock the single market, would be to *move away from attempts to achieve detailed harmonization* to a regime where each Member State recognized as valid the decision of another Member State's Courts in the criminal area with the minimum of formality' (Council of the European Union 1998 (emphasis added); Nilsson 2005). In this original view of mutual recognition, the fact that all the EU Member States are bound by the same international human rights instruments should suffice to justify establishing between them the mechanism of mutual recognition of judicial decisions adopted in criminal matters. And indeed, such has been hitherto the approach adopted by the Council of Europe instruments which promote mutual recognition on the criminal field: although these instruments, such as the 1970 European Convention on the International Validity of Criminal Judgments¹¹ or the 1972 European Convention on the Transfer of Proceedings in Criminal Matters,¹² contain certain safeguard clauses ensuring that criminal sanctions adopted by one state will not be enforced in another in violation of the latter's international obligations or of the 'fundamental principles of its legal system',¹³ they do not presuppose

that both states will have implemented principles such as respect for the rights of defence or the presumption of innocence, through similar or comparable national legislation.

Whatever the original intent behind the introduction of the concept of mutual recognition in the establishment of the AFSJ, it soon became clear that, far from rendering unnecessary the adoption of common standards, the mutual recognition of judicial decisions in criminal matters could in fact constitute an incentive for further harmonization, especially where the level of protection of certain fundamental rights in criminal proceedings varies among EU Member States. The idea of such complementarity between approximation of national laws and mutual recognition was recognized already by the Tampere European Council of 15–16 October 1999, which launched the idea of an AFSJ, and asked the Council of the EU and the Commission 'to adopt [...] a programme of measures to implement the principle of mutual recognition [including on] those aspects of procedural law on which common minimum standards are considered necessary in order to facilitate the application of the principle of mutual recognition, respecting the fundamental legal principles of Member States' (European Council 1999: par. 37).

Of course, whether mutual trust between EU Member States should be enhanced through harmonization measures—or whether, instead, mutual trust can be *presupposed*—cannot be dissociated from the fact that these states share a common *acquis* in the area of fundamental rights, in particular since they all are parties to the most important instruments of the Council of Europe. But this argument is not a decisive one. Council of Europe instruments do not cover all areas in which harmonization may be required in the EU in the field of criminal law. They often impose only *minimum* standards, particularly in the area of fundamental rights, and they therefore are an insufficient response to the risk of *divergences* between the Member States beyond those minimum standards, in the absence of any attempt at approximation under EU law.

In the establishment of the AFSJ, it is perhaps in the field of fundamental rights that the choice between harmonization on the one hand, and mutual recognition presupposing mutual trust on the other hand, presents itself in the purest form. And it is here, too, that the substitution of a logic of monitoring for the alternative between harmonization and mutual trust has been most heavily discussed: indeed, the scenario of human rights monitoring performed by the European Union on its Member States, in order to provide each Member State with the assurance that, if a serious threat to fundamental rights exists in another, this will be identified and reacted to as appropriate, has been explored in the period 2000–5 (De Schutter, 2008, 2009). This scenario was finally implemented in part only, essentially because of the fear that tasking the EU institutions with such a role would be competing with the kind of monitoring

performed by the Council of Europe bodies. However, the main weakness of fundamental rights monitoring within the European Union is neither that it is overambitious, nor that the Union would somehow exceed its mandate by developing into a 'human rights organisation' (Von Bogdandy 2000): it is rather that such monitoring is misdirected, because it is conceived as a top-down mechanism, aimed at verifying compliance with a predefined set of norms. The next part argues that monitoring the Member States is chiefly useful as a search device. Evaluation allows comparisons to be made. It allows each state to learn from the others. It therefore not only cements mutual trust, although this may be a desirable by-product, but first and foremost provides hope that, in the future, the choice between mutual recognition and harmonization will be better informed, and that, beyond those two branches of the classic alternative, mutual evaluation will emerge as a coordinating tool in its own right.

3 Evaluation mechanisms in the area of freedom, security, and justice: a typology

In order to understand the potential of such evaluation mechanisms in experimentalist governance architecture, we should first acknowledge the ambiguity of the position of the EU institutions concerning the criteria which they should follow when choosing, in the terms of the classic view, between mutual recognition and harmonization. For instance, the European Commission states in its July 2000 Communication on the Mutual Recognition of Judicial Decisions in Criminal Matters: 'Not always, but often, the concept of mutual recognition goes hand in hand with a certain degree of standardisation of the way states do things. Such standardisation indeed often makes it easier to accept results reached in another state. On the other hand, mutual recognition can to some degree make standardisation unnecessary' (European Commission 2000a: par. 3.1). The Hague Programme, too, remains vague on this crucial question. While this programme is intended to define the Union's agenda in the field of justice and home affairs for the years 2005–10, it simply mentions that the mutual trust on which mutual recognition of judicial decisions is based could be enhanced by a number of means, consisting of both legal measures and operational initiatives, and including in particular the 'progressive development of a European judicial culture based on diversity of the legal systems of the Member States and unity through European law'; 'a system providing for objective and impartial evaluation of the implementation of EU policies in the field of justice', providing 'the certainty that all European citizens have access to a judicial system meeting high standards of quality'; the 'development of equivalent standards for procedural rights in criminal proceedings, based on studies of the existing level of safeguards in Member

States and with due respect for their legal traditions'; 'the establishment of minimum rules concerning aspects of procedural law (...) in order to facilitate mutual recognition of judgments and judicial decisions and police and judicial cooperation in criminal matters having a cross-border dimension'; and finally, the approximation of substantive criminal law as regards 'serious crime with cross border dimensions', as provided by the EU Treaty (European Council 2004: par. 3.2 and 3.3).

This leaves to the European legislator an almost unlimited margin of appreciation. When it commented on The Hague Programme adopted by the European Council of 4-5 November 2004, the House of Lords urged caution on the question of approximation of the criminal laws of Member States in order to facilitate mutual recognition, emphasizing that 'this is an area where the principle of subsidiarity will come prominently into play and due observance of it will be necessary' (House of Lords Select Committee on the European Union 2005: par. 40). But what precisely the principle of subsidiarity might entail in this area remains unaddressed. The Protocol on the application of the principles of subsidiarity and proportionality, appended to the 1997 Treaty of Amsterdam, emphasized that subsidiarity is 'a dynamic concept', whose meaning will depend on the evolution of the circumstances: the principle of subsidiarity, it stated, 'allows Community action within the limits of its powers to be expanded where circumstances so require, and conversely, to be restricted or discontinued where it is no longer justified' (par. 3). It also imposed a requirement that any action of the Union subject to the principles of subsidiarity and proportionality be justified by reference to these principles (par. 4). Most importantly, it shed further light on the content of these requirements. The principle of subsidiarity requires that it be demonstrated that 'the objectives of the proposed action cannot be sufficiently achieved by Member States' action in the framework of their national constitutional systems and can therefore be better achieved by action on the part of the Community'. The verification of this condition may be influenced by considerations relating to the question whether 'the issue under consideration has transnational aspects which cannot be satisfactorily regulated by action by Member States'; whether 'actions by Member States alone or lack of Community action would conflict with the requirements of the Treaty (such as the need to correct distortion of competition or avoid disguised restrictions on trade or strengthen economic and social cohesion) or would otherwise significantly damage Member States' interests'; or whether 'action at Community level would produce clear benefits by reason of its scale or effects compared with action at the level of the Member States' (par. 5). Of these, the second justification for an intervention by the Union is clearly the most relevant in the establishment of the area of freedom, security, and justice: where the divergence between the Member States' approaches to a certain issue result in an obstacle to their mutual cooperation and, thus, threaten the aim of an area of freedom, security, and justice, this may call for

the approximation of national legislation, administrative regulations, or practices; in addition, according to the principle of proportionality, the intervention of the Union should be limited to what is necessary.

There are a number of signs indicating that mutual evaluation is emerging as a policy mode in its own right, either as a substitute for the mutual recognition/harmonization alternative or as a means to identify any divergences between the member States which may call for harmonization, consistent with the principle of subsidiarity which has just been recalled. A general monitoring, performed by mechanisms established within the European Union, of EU Member States' compliance with the values on which the Union is founded may never be established. But other, lower-profile forms of monitoring have recently been developing, in recognition of the need to ground mutual cooperation on a firm basis (Weyembergh and de Boliely 2006). Under one model, the European Commission is assigned a leading role, corresponding to its function under the EC Treaty as guardian of the Member States' obligations: it monitors the implementation of specific instruments adopted under title VI of the EU Treaty, on the basis of information collected from the national authorities. Under a second model, the Member States organize among themselves a form of peer evaluation, in order to improve the mutual understanding of one another's approaches to certain issues of common interest (such as the policing of external borders or the fight against terrorism), and to exercise political pressure on the Member States where certain deficiencies are identified. More recently, a third and more ambitious model, which may be seen as a synthesis an extension of these two existing models, has been proposed.

3.1 *The evaluation of the implementation of third pillar instruments*

A first category of 'evaluations' in fact aim, at a rather modest level, to compensate for the absence of infringement proceedings filed by the Commission against the Member States under Title VI of the Treaty on the European Union, in situations where they would fail to comply with their obligations under EU law, especially in the implementation of framework decisions. It has become typical for these instruments to require the Member States to report to the Commission, within a prescribed deadline after the period left for implementation has expired, about the implementation measures adopted; the Council is then expected to assess implementation with the framework decision on the basis of a report prepared by the Commission following the receipt of this information.¹⁴ Indeed, even in the absence of an explicit legislative mandate to that effect, the European Commission has occasionally considered that it should present such an evaluation of the implementation measures adopted by the Member States, putting forward the importance of the instruments concerned (European Commission 2004d: 3).

The impact of such evaluations is limited (de Biolley and Weyembergh 2006: 75–98). The information sent to the European Commission by the EU Member States relates to the implementation of a particular legislative instrument, rather than to the full set of measures adopted in a certain policy area (for more details, see European Commission 2001a: par. 1.2.2). These evaluations moreover are concerned only with the question whether the Member States have adopted the measures required under these instruments: they do not examine whether these implementation measures comply with the requirements of fundamental rights nor do they address whether, in the light of the difficulties encountered in the implementation phase, the legislative instruments adopted by the Union may have to be amended, or even completely redesigned. They concentrate on the adoption of legal measures by the Member States: they are silent about the practical impact of such measures, and about the question whether these measures effectively contribute to the establishment of the AFSJ. Although there have been attempts to move beyond the practice of evaluations based exclusively on the legal measures adopted, in particular, as regards the implementation of the Council Framework Decision of 13 June 2002 on the European arrest warrant and the surrender procedures between Member States,¹⁵ the uneven quality of the information on which this was based has been recognized as a serious deficiency of the process (European Commission 2006c). The Commission did suggest farther-reaching evaluation mechanisms, such as involving independent experts in monitoring the effective compliance by all the EU Member States with the fair trial requirements imposed under the proposed Framework Decision on certain procedural rights in criminal proceedings throughout the European Union, or imposing on the national authorities an obligation to collect statistics about the impact of this instrument (European Commission 2004d). But these proposals for strengthening the monitoring of implementation of EU law apparently met with strong resistance.

In sum, the evaluations which have been set up under specific instruments adopted under the third pillar hitherto have served to monitor the compliance of the Member States with their obligations, in a classical top-down fashion; but they are not seen as a potential source of reflexivity for the EU institutions or as providing an opportunity for collective learning between the Member States. It is as if the European legislator could do no wrong. And it is as if the adoption of legal measures, by itself, would be sufficient to create the conditions which will ensure that they will achieve their objective, however diverse and evolving the settings in which these measures are to be implemented.

3.2 Peer evaluation

In the kind of evaluation discussed above, the European Commission plays the central role as the guardian of the Member States' obligations. In contrast, peer evaluations have developed in certain areas. One of the oldest and most

interesting forms of peer review organized in the justice and home affairs field results from the establishment of a Standing Committee on the evaluation and implementation of Schengen, entrusted with evaluating both the degree of preparedness of the states who are candidates for participation in the Schengen Convention, and the level of compliance of existing signatory states.¹⁶ This committee is composed of one high-ranking representative from each signatory state. Its delegations, composed of inspectors representing the Member States willing to contribute (each state funding its own representative within the group), visit the countries subject to the evaluation procedure, according to a work programme defined initially by the Executive Committee, and now by the Council of the EU. The mutual evaluation is organized on the basis of the information collected through these visits as well as information provided by the host state. The purpose of this monitoring is not only to evaluate whether all the preconditions for applying the Convention Implementing the Schengen Agreement in a candidate state have been fulfilled but also to 'seek solutions to the problems detected and [to] make proposals for the satisfactory and optimal implementation of the Convention' in the existing signatory states.

The peer review mechanism is based on inspections in the states concerned as well as on a written questionnaire-based procedure.¹⁷ This procedure, which remains fully confidential, leads to the adoption of political conclusions by the Council of the EU, which may approve recommendations adopted by the Working Group. After a state has been subjected to an evaluation, it must present a follow-up report stating how it met the recommendations made by the experts. The follow-up may identify the measures which were adopted in response to those recommendations; or it may explain why certain reforms could not be implemented immediately, for example, because they require the reinforcement of the existing capacities, for which the necessary budgets may be lacking; in certain cases, the states concerned have contested the recommendations addressed to them.

The evaluation mechanism is thus conceived to allow the detection of any problems encountered in the implementation of the Schengen Convention, and to identify solutions proposed for applying the Convention in the most satisfactory and effective manner. The participating states are thus placed under a close supervision, focused not only on the legal transposition of the Schengen acquis but also—and primarily—on its practical implementation, in areas such as border controls, visas, protection of personal data, or the expulsion or readmission of foreign nationals. In the development of the Schengen evaluation mechanism, a number of remarkable evolutions have taken place.¹⁸ In particular, the catalogues prepared in order to facilitate the implementation of the Schengen acquis—compendia of best practices in areas such as, for instance, the crossing of external borders and the delivery of visas¹⁹—have taken into account in the practice of evaluations, even though these

catalogues have no binding legal effect. The reports of the inspection visits have also been structured in a more harmonized way, thus ensuring the possibility of comparison between evaluations.

Peer assessments broadly similar to the Schengen evaluation mechanism have developed in the fields of terrorism and of organized crime. The Joint Action of 5 December 1997 establishing a mechanism for evaluating the application and implementation at national level of international undertakings in the fight against organized crime²⁰ provides perhaps the best illustration. This peer evaluation, intended to cover at least five states per year, is supervised by the members of the Multidisciplinary Working Party on Organized Crime (MDW). Acting on the proposal of the presidency of the Council, the MDW defines the specific subject of the evaluation²¹ as well as the order in which Member States are to be evaluated. The evaluation teams comprise three experts for each Member State subject to the evaluation drawn from a list of experts presented by the other Member States, and include in addition one or two members of the General Secretariat of the Council, one representative of the Commission, and occasionally a member of a body such as (depending on the subject of the evaluation) Europol or Eurojust. On the basis of the answers of the Member State concerned to a questionnaire and of a visit in that Member State allowing the evaluation team to meet the officials involved, a draft report is prepared, which is transmitted to the MDW along with the comments of the state which were not accepted by the evaluation team. The MDW adopts conclusions by consensus, following a presentation of their report by the evaluation team, and the explanations received from the State subject to the evaluation. Those conclusions are transmitted to the Council, which may address recommendations to the Member State concerned and invite it to report back to the Council on the progress it has made by a specific deadline. The follow-up of the recommendations is generally weak: although most states do respond to the recommendations addressed to them, the information they send to the Council does not lead to any further discussions. Interestingly however, the MDW has occasionally included recommendations addressed to the Council itself, or to Europol: this suggests that, although conceived initially for the monitoring of the Member States' application and implementation of international undertakings in the field of organized crime, this mechanism has the potential to bring about improvements also in the approach developed at European level in this field. On the other hand, the mechanism does not contribute to the accountability of the executives towards either the national parliaments or to civil society organizations, since the whole process is in principle confidential, although the Member States evaluated may if they wish make public the reports under their own responsibility.

This system was further built upon in order to ensure a form of peer evaluation of the action of the Member States against the threat of terrorism.²²

The mechanism is placed under the supervision of the 'Article 36 Committee', a Committee of high-level national public servants established under Article 36 EU in order to prepare the discussions within COREPER and the Council in the fields of police cooperation and judicial criminal cooperation. For each cycle of evaluation in this area, this Committee chooses one theme.²³ Within six weeks after receiving the reply of the Member States to the questionnaire prepared on that theme by the Presidency of the Council, an evaluation team composed of two national experts from other Member States and assisted by the General Secretariat of the Council and the Commission may if appropriate travel to that Member State, in order to clarify the replies to the questionnaire: a programme of visits is arranged to that effect by the Member State visited on the basis of the evaluation team's proposal, for interviews with the political, administrative, police, customs and judicial authorities, and any other relevant body. The members of the Article 36 Committee receive the draft report of the evaluation team, along with any comments of the state concerned which the evaluation team did not wish to include. On the basis of a discussion introduced by the presentation of their report by the members of the evaluation team, the Article 36 Committee adopts conclusions by consensus. At the end of a complete evaluation exercise, the Council is informed of the results of the evaluation, and it may address recommendations to the Member State concerned and invite it to report back to the Council on the progress it has made by a certain deadline.²⁴ The information collected by the evaluation teams in this process, as well as the country-specific recommendations, are confidential. Only the synthesis reports adopted by the Council at the close of an evaluation cycle are public, and are transmitted to the European Parliament; however they contain no references to specific states.

3.3 Strategic evaluations

Each of the evaluation processes described above presents a number of deficiencies. The evaluations by the Commission of the implementation of certain specific instruments adopted under Title VI of the EU Treaty essentially focus on the adoption of legal measures by the Member States, rather than on the practical effectiveness of the policies to which those instruments seek to contribute. If they add to our understanding of the adequacy of those instruments themselves, this results from chance rather than from design. As a tool to improve the reflexivity of European policies, which should allow for those policies to be revised in the light of their impact in different settings, they are poor. And even as a tool to exert pressure on Member States in order to ensure that they adopt all the implementation measures required, these meet with only partial success, since the Commission cannot file infringement proceedings for failure to comply with the obligations imposed by instruments adopted under the third pillar of the EU Treaty.

In contrast, the peer evaluations conducted in order to contribute to the implementation of the Schengen *acquis*, or in the fields of organized crime or terrorism, have a potential to bring about policy changes in certain Member States, and may contribute to mutual learning in certain fields where the Member States have adopted significantly different approaches. But their contribution to improving the accountability of the governmental departments concerned is limited: with few exceptions, their results are not public, and any pressure exercised on a state by the other Member States within the Council of the EU cannot be relayed by national parliaments or by civil society organizations. In addition, as clearly illustrated by the preparation of compendia of best practices such as the Schengen catalogues in the Schengen evaluation mechanism, these evaluations presuppose that, for any question of common interest, there exists one 'adequate' or 'best' way to implement certain predefined objectives: while the same processes may also occasionally lead to 'discover' new approaches to old problems, on the basis of certain experiments conducted by one Member State, this is not the explicit aim of the peer evaluations—and even where it happens, the end goal still appears to be greater uniformity, even if this may take the form of the adoption by all the states of certain best practices identified in one of their number. Finally, these evaluations, even considered together rather than individually, remain fragmentary and *ad hoc*, rather than guided by any overarching vision about how evaluation may contribute to the rationality and reflexivity of EU policies.

This may be changing. The Lisbon Treaty provides that the Council may adopt measures 'laying down the arrangements whereby Member States, in collaboration with the Commission, conduct objective and impartial evaluation of the implementation of the Union policies referred to in [Title IV: Area of Freedom, Security and Justice] by Member States' authorities, in particular in order to facilitate full application of the principle of mutual recognition. The European Parliament and national Parliaments shall be informed of the content and results of the evaluation'.²⁵ Clearly, the intention is not solely to supervise the faithful implementation by the Member States of instruments adopted in this field (something which, under the new framework established by the Treaty of Lisbon, should in any event be facilitated by recourse to the more classical means currently used in Community law). Rather, that evaluation should serve other goals, primarily to establish mutual trust between the Member States' national authorities, and to ensure that the development of EU policies are fully informed by the difficulties encountered in practice by the Member States in the course of the implementation of these policies, in order to allow for them to be revised in the light of such obstacles. In addition, not only the European Parliament but also national parliaments are to be involved in the evaluation. This should increase the pressure on the Member States, and it should improve the accountability of the Executives who fail to comply with their obligations under EU Law. But it should also contribute to each national

parliament gaining a better understanding of the stakes of European integration, and of the nature of the obstacles faced in other Member States in the implementation of European policies developed in the field of freedom, security, and justice.

Following upon The Hague Programme adopted by the European Council on 4 November 2004 (European Council 2004: par. 3.2), the Council and the Commission adopted on 2–3 June 2005 an Action Plan listing the setting up of a system for objective and impartial evaluation of the implementation of EU measures in the field of Freedom, Security, and Justice as the first priority.²⁶ The result was the presentation by the Commission, on 28 June 2006, of a communication on the evaluation of EU policies on Freedom, Security, and Justice (European Commission 2006d). This communication goes much further, however, than what the Member States had anticipated. It proposes a form of systematic evaluation of the implementation of the EU policies in these fields,²⁷ aimed not only at ensuring compliance with the Member States' obligations under EU law but also at evaluating EU policies as such. The Hague Programme had stated that '[e]valuation of the implementation as well as of the effects of all measures is (...) essential to the effectiveness of Union action' (emphasis added). In line with this mandate, the communication on strategic evaluations deliberately seeks to ensure that the evaluation of developments at Member State level will serve to improve the design of EU policies. These policies are therefore being 'tested' at the same time that the Member States' implementation is being 'monitored': the evaluation of the implementation of the EU policies ensures a feedback on the latter themselves, which may have to be revised in the light of the problems encountered in their implementation or the—perhaps unintended—impacts they produce; and, beyond the aim of monitoring as a means to ensure compliance, evaluation serves the aim of promoting learning, by the comparisons it should allow of the experiences of the different Member States in the implementation process.

The mechanism proposed in the communication consists in the Member States providing the Commission with information about the implementation of EU policies in the fields of freedom, security, and justice, by the regular delivery of 'factsheets' (one for each policy area), describing the achievements of each Member States on the basis of a relevant set of indicators. Such factsheets should be communicated twice every five years, since they will focus on 'slow-moving outputs and results and on medium-term data' (European Commission 2006d: par. 35). The information contained in these factsheets would be commented upon by the relevant stakeholders. The Commission would then prepare an 'evaluation report', including certain political recommendations. Finally, where justified, an 'in-depth evaluation report' would be prepared by the Commission in specific areas. 'Strategic' evaluations thus conceived should add value to the current practices as described above, according to the Commission, notably by (European Commission 2006d: par. 33).

- (a) focusing on *policies* (or coherent subsets), rather than individual instruments (for instance, evaluation of the common immigration policy);
- (b) analysing the *coherence* of different instruments within a given policy (e.g. how financial programmes support and facilitate implementation of the EU legislation in a given field);
- (c) investigating how a certain policy contributes to the *overall objective* of establishing an Area of Freedom, Security, and Justice;
- (d) determining the *overall rate of achievement* of that general objective; and
- (e) assessing achievement of an overarching objective in the field of freedom, security, and justice (for instance, safeguarding of fundamental rights).

This goes far beyond a banal practice of monitoring Member States' compliance with their legal obligations under the EU Treaty, in particular, to compensate for the absence of infringement proceedings filed by the Commission under Title VI EU. Rather, the strategic evaluation of the EU policies adopted in the field of freedom, justice, and security constitutes a means to ensure that (a) beyond the compliance achieved by the adoption of legal instruments, the effectiveness of the measures adopted (their ability to achieve the objectives of the policies pursued) is measured (b) the political objectives are regularly redefined in the light of the lessons which may be drawn from implementation by the Member States; and (c) our common understanding of the requirements of an 'area of freedom, security and justice' is progressively transformed by this iterative process. In sum, this 'strategic evaluation' should serve to transform what currently may be seen as a liability (the absence between the Member States of a common understanding of the end-goal of an AFSJ, combined with the inability of the Commission to effectively impose on them such an understanding) into a virtue: the very ambiguities about the meaning of this project are productive, in that they allow a collective learning to take place in which the 'principals' learn from the 'agents' and, in the light of the information provided by the agents, may be led to revise their understanding of what political initiatives may be required.

As confirmed by the mostly sceptical reactions it triggered from the Member States,²⁸ the communication on strategic evaluations goes beyond what is envisaged in the Lisbon Treaty. It envisages the evaluation process leading to an improved accountability of policymakers, both at European and at national level, through the dissemination of the results of the evaluation and the involvement of a variety of actors in the process: in particular, the 'evaluation report' prepared by the Commission on the basis of the factsheets communicated by the Member States will be transmitted not only to the Council and European Parliament but also to the European Economic and Social Committee and the Committee of the Regions; it will be 'disseminated as appropriate

to wider audiences, including via ad-hoc public events' (European Commission 2006d: par. 21); more generally, the Commission will 'ensure that the views of the civil society will be taken into account and will establish appropriate mechanisms to ensure its participation in the evaluation of all policies in the area of freedom, security and justice' (European Commission 2006d: par. 16).

The form of 'strategic evaluation' proposed should not be seen as a form of Open Method of Coordination in the canonical definition given the latter by the Lisbon European Council of March 2000. The Member States are not requested here to prepare action plans which will be subjected to peer review, and lead to the adoption of guidelines by the Council: they are, rather, to provide factual information to the Commission about the effectiveness of the policies they are pursuing, for the latter to draw political conclusions and stimulate debate about the need to revise EU policies adopted in the field.²⁹ Nevertheless, the dimensions of mutual learning and of peer review are not absent from the strategic evaluations. Even more importantly, these strategic evaluations are devised as a response to the uncertainty we face in the fields they will cover: the reason why there is a need to evaluate the effectiveness of the EU policies developed in these fields is that 'good' answers to the questions of how to create the mutual trust required for the mutual recognition of judicial decisions in civil and criminal matters, how to effectively prevent organized crime, or how to combat illegal immigration—to mention only those examples—are not readily available, and that the initiatives adopted so far may appear to be based on misguided information, not to have anticipated certain secondary effects, or to have underestimated certain obstacles to implementation by the national authorities. Indeed, not only are the means to be permanently 'tested' in the light of the national authorities' experience with the implementation of EU policies, the ends themselves—*what we mean* by the establishment of an area of freedom, security and justice—need to be redefined, or reinvented, as we unpack the implications of seeking to implement them.

4 The potential of evaluation

On their surface, the strategic evaluations the European Commission proposes to introduce in the field of freedom, security, and justice, should provide it and the Member States the information they require to improve EU policies in the six areas they will cover. The practice of such evaluations fits into the broader framework of improving governance in the European Union: in the July 2001 White Paper on Governance, the Commission had already emphasized the need for 'a stronger culture of evaluation and feedback (...) in order to learn from the successes and mistakes of the past' (European Commission 2001b: 22). And it constitutes a clear recognition that, such an evaluation is not

satisfactorily organized in the standard inter-institutional division of tasks in EU law and policy-making. Neither the European Commission nor the European Parliament has all the information required from the Member States to perform such evaluations; indeed, to make this information available in a transparent and non-selective manner is precisely what the June 2006 communication seeks to achieve. The Council of the European Union has been developing a practice of peer assessment for almost ten years in certain well-defined areas, but it is seriously handicapped by Member States' natural tendency not to put excessive pressure on one other, especially in fields such as law enforcement, which are traditionally associated with the core of national sovereignty. In addition, it is difficult for the Council to question its general orientations in the light of possible resistance in certain Member States, since this would risk undermining its credibility and encouraging non-compliance. Indeed, it is perhaps at this last level that the novelty of the 'strategic evaluations' proposed by the European Commission is most striking: rather than offering to monitor Member States' compliance through certain instruments or predefined policy options, the strategic evaluations explicitly consider that difficulties in the implementation phase may indicate not that the Member States concerned are acting in bad faith, or are unwilling to contribute to the common objective—but that these predefined instruments or options may be misconceived, or may have underestimated the obstacles resulting from the need to apply them in particular settings whose dynamics could not be anticipated.

Yet, as conceived in the 2006 communication, the reflexive potential of evaluation may be lost, if a number of conditions are not fulfilled. By reviewing the aims of evaluation in the fields covered by the communication, we may hope to shed some light on the conditions which should be created for such evaluations to effectively contribute to the legitimacy and efficiency of these policies. However, as we will discover, the relationship between the different objectives of the system of strategic evaluation proposed remains ambiguous and, if not considered in its own right, could become a source of tension. Five objectives at least may be distinguished. The two first objectives, which only a thin line separates from one another, are considered together.

4.1 Monitoring the quality of implementation of EU policies and ensuring feedback on them

Article 61C of the Treaty on the Functioning of the European Union mentions the need to establish 'an objective and impartial evaluation of the implementation of the Union policies [in the area of freedom, security and justice] by Member States' authorities'. This refers to the aim of monitoring whether or not the EU Member States loyally cooperate in the implementation of these policies, not only by transposing the instruments which are adopted but also,

for instance, by ensuring that their authorities cooperate with those of other Member States, or that the operational measures required for the implementation of EU instruments are taken. But the 2006 communication mentions a quite different aim, which is to 'improve policy-making, by promoting systematic feedback of evaluation results into the decision-making process' (European Commission 2006d: par. 7). Here, EU policies themselves, rather than their implementation by the Member States, come into question: from an evaluation of the implementation measures by the Member States, the communication shifts to an evaluation of the effects of such implementation, in order to improve the policies adopted at EU level. The notion of evaluation on which this shift relies is the one defined initially in the context of the reform of EU governance, at a time when both the legitimacy and the efficiency of the Union's policies were under heavy criticism. Among the many initiatives which this reform has led to since 1999–2000 is the adoption of an internal communication on evaluation which states that

Evaluation is '*judgement of interventions according to their results, impacts and the needs they aim to satisfy*'. It is a process in which DGs and Services engage in order to identify what can be learned for policy and planning. Furthermore, evaluation findings should contribute to Commission level decision-making on priorities and resource allocation.³⁰

The 2006 communication alludes to this where it writes that the evaluation mechanism proposed 'is based on this comprehensive definition which, in the Commission's view, should allow a full understanding of the quantity and quality of results achieved on freedom, security and justice' (European Commission 2006d: par. 7). It will be noted, however, that a distinction may be made between policy feedback and policy learning. As explained by Anton Hemerijck and Jelle Visser: 'Policy learning is analytically distinct from policy feedback in that it essentially gives pride of place to the reflexive and evaluative, both cognitive and normative, activities of policy actors' (Hemerijck and Visser 2006: 37). In Kuhnian terms, one might say that policy learning seeks to question the policy paradigm itself, and not only the adequacy of the implementation measures adopted under the paradigm guiding the policymakers in a particular policy area. If we use this distinction in that sense, although the 'strategic evaluations' proposed by the Commission might lead to policy learning within the Member States themselves (a question which is further examined below), it is more doubtful whether it will ensure genuine learning in the design of the policies at EU level, whereas the evaluation is designed to judge interventions according to their results and impacts, in the light of the needs they aim to satisfy, these needs themselves—the general objectives, typically set by the European Council—will presumably not be questioned in this process.

The definition of evaluation quoted above is too narrow in another respect. If we take this formulation literally, this evaluation should consist in judging

the impact of interventions. However, one of the main aims of strategic evaluations should also be to judge the impact of the *absence* of interventions, that is, of the *failure* of the European Union to harmonize national laws, regulations, and practices, or to improve the coordination, through any alternative means, between national authorities. These evaluations are explicitly stated to focus on *policies* (such as, for instance, the common immigration policy) rather than on specific *instruments*. This creates the possibility that the information collected from the national authorities will highlight the need for more EU intervention, for instance, in order to encourage the diffusion of the best practices identified in one Member State or in order to ensure that certain measures adopted in one Member State (say, massive regularization of foreigners illegally staying on the territory) are not undercutting the efforts of another Member State in the same area (such as to discourage candidates to illegal immigration in the EU). The information collected from the national authorities are conceived as relating not only to the existing EU instruments, but more broadly to policy objectives identified at the European level.

Because the evaluations will cover measures adopted in areas where the European Union has not acted (or has not acted yet), they have the potential of both depoliticizing and repoliticizing the interpretation of the principles of subsidiarity and proportionality, which should guide the exercise of EU competences in the fields which it shares with the Member States. Indeed, by ensuring that the Commission and the Council will be informed of the full set of measures adopted by each Member State in a particular area, the answer to the question of whether the intervention of the EU would have a truly added value—insofar as the objectives cannot be sufficiently achieved by Member States' individual actions and can therefore be better achieved by action on the part of the Union, to paraphrase the treaties—will be based on evidence, and on the comparison of data from all the EU member States, rather than on the basis of mere intuition or on considerations relating to the political feasibility of any particular initiative. In that sense, a system of objective and reliable evaluation of the Member States' policies in the fields of freedom, security, and justice, should better insulate decisions about the desirability of EU intervention from political pressure: hence, the depoliticization of subsidiarity and proportionality this might entail. But at the same time, these principles would be repoliticized, insofar as the evaluations may be a tool for ensuring the participation of the European Parliament, the national parliaments and a wide range of other stakeholders in the discussion about which lessons should be drawn from the information pooled (European Commission 2006d: par. 11–16). These are not conflicting tendencies. They both point towards ensuring that agenda setting in the European Union and the sequencing of EU interventions are made more transparent and the subject of explicit deliberation, based on sound and comparable evidence concerning the evolution of policies developed at Member State level.

4.2 Promoting mutual learning

A third, and again distinct, aim of evaluation is to promote mutual learning between the Member States. Under an evaluation emphasizing the first aim identified above (that of ensuring compliance with certain predefined instruments or policies), uniformity (or at least convergence) between the Member States is seen as positive and desirable; and diversity, instead, is considered with suspicion. In contrast, where the focus is on mutual learning, diversity is cherished as a potential source of progress. The Member States are not encouraged to demonstrate that they act according to a script prepared for them; they are asked what original approaches they have to offer which might lead others to revise their own presuppositions about the most efficient approach. It is in the fulfilment of this aim that peer evaluations have an unparalleled potential, especially when it is conducted—as in the Schengen evaluation mechanism or in the mechanisms established in the areas of organized crime or counter-terrorism—by the counterparts, in the other Member States, of the very officials who are in charge of implementing a particular policy and who may be visited by an evaluation team. This kind of interaction between national civil servants may lead to blurring the differences between the respective positions of the 'evaluators' and the 'evaluated'; instead of the former controlling whether the latter effectively comply with what is expected by their European partners, the national agencies who are subjected to the evaluation may be developing original approaches towards certain problems faced also in other States, from which the evaluators might seek inspiration and may even wish to promote. The claim is not that such an identification and diffusion of best practices takes place effectively under the peer evaluations which are currently practiced in the fields of freedom, security, and justice—although it is more likely than not that examples of this could be found. Rather, the claim is that if mutual learning is one of the objectives of evaluation processes, peer evaluations may be the most adequate tool through which this can be achieved.

The enumeration of mutual learning among the aims of evaluation assumes that policy changes may develop not only incrementally, as a result of small-scale corrections to the dominant approaches in place through feedback mechanisms and as a result of trial-and-error processes, but also through cognitive or normative shifts in the policy-makers' understanding of causality chains or in the values guiding policy, that is, in the definition of the ends they seek to pursue. It assumes, further, that such shifts may result from the confrontation of policy-makers with other perspectives, or approaches, adopted in other Member States, towards the same problem. Certain conditions must be created, however, before such mutual learning effectively occurs, and in order that it may be successful. One set of conditions concerns the circumstances surrounding the learning process. For instance, a sense of crisis—policy-makers' conviction of that things cannot continue as they have previously and that the

perpetuation of routines is not a viable option—may enhance their willingness to learn, and thus create the necessary motivation to borrow from solutions developed elsewhere. In that sense, although not necessarily a condition for mutual learning, crises provoked by the failure of previous policies may facilitate policy changes.

Although some conditions favourable to learning cannot necessarily be created, others can. Thus, it may be presumed that if a particular experiment conducted in another jurisdiction is shared with a wide variety of actors in the 'receiving' jurisdiction, this will have greater chances of influencing policy debate and, perhaps, of bringing about changes. Similarly, if there exists, within the 'receiving' jurisdiction, an agency specifically dedicated to the understanding of such foreign experiments and to assessing whether the transposition of such experiments would be desirable, this could greatly contribute to overcoming bureaucratic inertia and the resistance of policy-makers who, in the face of uncertainty about whether change will be rewarding, might otherwise prefer to opt for the perpetuation of routines—for choice without search. In that sense, the reception structures may be more or less favourable to mutual learning: the wide diffusion of foreign policy experiments to a broad range of actors, as well as the establishment of expert bodies or think tanks whose mission it is to draw the attention of policy-makers to the need to explore those solutions, could greatly contribute to the success of mutual learning as one possible result of evaluation.

Another set of conditions relate to the channels of mutual learning—the process through which learning occurs or not. In particular, a contextualization both of the solutions developed in other settings and of the problems encountered in the 'receiving' jurisdiction seems necessary for learning to be successful. Solutions developed elsewhere cannot simply be presumed to be transposable to any other context: instead, what makes one approach successful in any particular situation will depend on a full range of factors which may or may not be present in the context in which that solution is being replicated. In what may be seen as one version of the 'garbage can' logic of decision making (Cohen et al. 1972), the available solutions risk predetermining the understanding of the problem to be addressed, rather than the problem being diagnosed independently of which solutions offer themselves. Therefore, any attempt by a 'receiving' jurisdiction to borrow from solutions developed elsewhere to similar policy problems should be preceded by an attempt to identify the conditions which allowed those particular solutions to be effective where they were first introduced, and by a diagnosis of the reasons why the approaches currently in place in the receiving jurisdiction have failed, which should be conducted independently of the existing catalogue of alternative policies. While foreign experiences may shed light on certain problems in the 'receiving' jurisdiction which might otherwise have been underestimated or ignored, they should not be seen as a substitute for

the analysis of those problems under the specific circumstances in which they have arisen. The risk of such 'decontextualized learning', in which solutions are prescribed irrespective of local conditions (Hemerijck and Visser 2006: 42), is especially high where the analysis of policy options is sectorialized, that is, where this analysis focuses on discrete areas of public policy, defined relatively narrowly, and thus detached from the analysis of the background conditions which may play a role in the success or failure of the policy options which are experimented. It seems contestable, for instance, to evaluate the policy of the Member States in the area of trafficking of human beings without considering different approaches to prostitution; or to evaluate their respective counter-terrorism strategies in isolation from the tools they develop to integrate third-country migrants residing on their territory and ethnic or religious minorities which may be tempted by violent radicalization.

4.3 *Enhancing mutual trust*

The Hague Programme adopted by the European Council mentioned that the mutual trust on which mutual recognition of judicial decisions was based could be enhanced by 'a system providing for objective and impartial evaluation of the implementation of EU policies in the field of justice' (European Council 2004: par. 3.2. and 3.3.). Indeed, a practice of evaluation may limit the risk of misunderstandings occurring between national authorities of different states, which may result simply from the differences between the legal systems in which they operate: evaluation thus conceived may be a means of ensuring that, however important those differences may seem, all the States at least comply with certain standards; and it may encourage a better knowledge of one another's system, facilitating in turn cooperation between the authorities concerned. But in an evaluation conducted for the purpose of creating mutual trust conformity will be rewarded: if there are differences, these will be minimized; rather than being an asset, original solutions to common problems are a threat, since they risk undermining mutual confidence.

4.4 *Stimulating democratic deliberation*

The 2001 White Paper on European Governance lists both participation and accountability—along with openness, effectiveness, and coherence—among the five principles of good governance (European Commission 2001b: 10). Evaluation of course, contributes to the effectiveness of EU policies. But it may also stimulate democratic debate and promote accountability. Provided with the results of an evaluation of the achievements of the Member State concerned in a particular policy area, opposition political parties, civil society

organizations, the media, and the public at large, not only will be better equipped to request explanations from decision-makers, and to critically gauge the justifications offered for pursuing particular policy options, they will also be more motivated to invest in any participatory mechanisms proposed to them. The broad discussion of alternatives to the dominant solution may provide public officials with an incentive to revise their routines. This may compensate at least partly for their fear that, by exploring those alternatives, they will betray established expectations and threaten acquired positions, which may be costly in electoral terms—especially since, as noted by March and Olsen, voters tend to sanction mistakes, more than the failure to explore untested opportunities, leading policy-makers to be generally risk-averse.³¹

Indeed, this may be one possible result of the strategic evaluations as conceived by the Commission (European Commission 2006a). In its 2006 communication, the Commission proposes that the 'factsheets' to be filled in by the Member States should be 'put out to consultation with relevant stakeholders and civil society' (European Commission 2006a: par. 24). It is indeed essential that the information provided by the Member States be completed from other sources. Although the strategic evaluation mechanism is ostensibly designed to improve EU policies rather than to monitor the contribution of each Member State to their implementation, the quality of the latter will also necessarily figure in these evaluations. National authorities therefore may be reluctant to provide the European Commission—and, thus, the other Member States and the broader public—with information which casts them in an unfavourable light. The 'factsheets' as completed by the national authorities therefore should be verified against any other information available, especially that collected by independent experts or civil society organizations. An interesting precedent in this regard is the establishment, by DG Employment, Social Affairs and Equal Opportunities of the European Commission, of networks of independent experts in order to monitor the implementation of the directives adopted in the equality field.³² These experts provide the Commission with crucial information not only on the legal measures adopted in each Member State, in a format which is generally more complete and systematic than what could be expected from national administrations, but also on any gaps which those measures present, or the problems met in their practical implementation. Such a mechanism could equally be conceived in other areas, in order to ensure that the evaluation is based on reliable and balanced reports.³³ But even in the absence of such independent assessment, the publicity given to the factsheets filed by the Member States, combined with the possibility for other interested parties, including non-governmental organisations, to complement or contradict this information—and with the possibility for opposition political parties in national parliaments to hold the government accountable for the information it provides—should ensure its trustworthiness.

4.5 Combining the diverse aims of evaluation

While all these aims potentially served by an evaluation mechanism are clearly desirable, it is a distinct question whether they all can be pursued at the same time, through similarly conceived processes. Two views of this are possible. In one view, we would have to choose between two models, neither of which is capable of fulfilling all the aims listed above. The first model is of relatively closed, peer-review mechanisms, through which Member States may have frank exchanges on the basis of information which—because it is not public—can be presumed to be more reliable than if it were to be shared and therefore potentially used against them. This model, directly involving civil servants from the Member States in the evaluations, would also be more conducive to mutual learning, since the evaluations are conducted by the very individuals who could benefit most from it. The second, alternative model is put forward by the 2006 communication on the evaluation of EU policies in freedom, justice, and security. The communication intends the strategic evaluation mechanisms it proposes to be broadly participatory, with an involvement of a wide range of actors including civil society organizations; and it envisages that the factsheets prepared by the Member States will be made public. As already noted, this approach—in line with the White Paper on European Governance of 2001 (European Commission 2001b)—has the potential to enrich democratic deliberation and to improve the accountability of policy-makers, by obliging them to provide justifications for not exploring certain alternatives to the prevailing routines. But such openness, it could be argued, may also constitute a disincentive for the disclosure of certain failures or resistance encountered in the implementation of policies: it may be difficult to convince national authorities to be fully transparent about such failures or the nature of the obstacles they are facing, since this not only may be held against them in internal electoral debates but could also provide a pretext for proposing further interventions by the EU in areas where the national authorities appear particularly jealous to preserve their national sovereignty.

But this is not a true dilemma. In fact, the virtues attributed to peer-review evaluation mechanisms can also result from more openness and transparency, rather than less. As we have seen, the best way to ensure that policy learning takes place may be to provide a broad range of actors with an incentive to challenge dominant policy paradigms in the light of the available alternatives, and a wide discussion about such alternatives might make them more attractive to decision-makers. Hence although peer-review mechanisms, left entirely in the hands of the States, may be conducive to mutual learning thanks to the direct exchanges they permit, the resulting advantages may be more than offset by the lack of involvement of other actors, especially at the national level, whose support may be decisive for any policy learning which occurs to lead to actual improvements in the design and implementation of policies.

As to the reliability of the information provided by the Member States, although it may to a certain extent be achieved by this information not being made public, this also may result from any information provided by the national authorities being widely discussed and cross-examined, by national parliaments and civil society organizations. In sum, while the secrecy of peer-review mechanisms may seem to present certain advantages, these can be obtained through an entirely different strategy, which emphasizes openness over closure and publicity over confidentiality.

A more serious dilemma may be between the prescriptive and the non-prescriptive dimensions of evaluation—in other words, between monitoring and learning in evaluation. Where monitoring compliance with predefined instruments or policies is emphasized, the Member States will have a natural tendency to present their practices and results in the best possible light; where the focus is, instead, on collectively deciding what approaches should be privileged, or on evaluating whether the EU policies work, they may become more open about their failure to achieve results and the obstacles they face. In a form of evaluation promoting mutual learning, Member States will explain how they have developed different approaches; in one which seeks to monitor compliance with agreed upon instruments or objectives, or which is seen as a contribution to creating mutual trust, they will dismiss these differences as merely superficial, and seek to convince their interlocutors, instead, that what they are doing is really the same as what others are doing—or that it better follows the agreed script.

In the face of this second dilemma, it is again tempting to contrast, in a binary mode, two forms of 'evaluation': one geared towards verification of compliance with commonly agreed objectives, and which could lead to addressing recommendations to the states concerned when they deviate from those prescribed objectives, and the other aimed at 'exchange of experiences', without any monitoring dimension. But the practice of the institutions illustrates the fragility of this distinction. Both in the fields of asylum and immigration³⁴ and in the broader area of freedom, security, and justice, the exchange and pooling of information between the Member States, coordinated by the Commission, has been preferred to the adoption of guidelines by the Council and the preparation of national plans to implement them. But this shift has been tactical rather than strategic, and its consequences should not be overemphasized: once it agrees to report on its policies in a particular area, and to have this information made public, each Member State accepts to be held accountable in principle, both to one another (for any externalities of its unilateral actions) and to public opinion at home and abroad (for any gap between its own achievements in reaching a goal collectively recognized as desirable and the achievements of its neighbours confronted with a similar set of circumstances). Whether evaluation takes the form of OMC-like processes or of less demanding exchanges and information pooling, some form of

diagnostic monitoring takes place, in which the achievements of each Member State are related both to the conditions it faces and to the policies in place, in order to assess whether the policies should be changed or, instead, inspire others, thus breaking down the distinction between learning and monitoring. The difference however, is in the mode of identification of the best means to achieve the common objectives: whereas, in an OMC-like process, what the best practices are is the subject of a deliberation between the Member States within the Council, in procedures limited to mutual information, these best practices are, at best, progressively defined by the Commission in the reports it prepares on the basis of the information coming from the Member States; at worst, this identification remains implicit, it is never openly discussed for its own sake, and as a result, what is most desirable may be understood quite differently by each actor. In defence of this approach, it may be said that such an ambiguity may be productive: it may encourage Member States to explore a diversity of approaches, experimenting in ways from which the others may then seek inspiration, and replicate, partially or wholly, taking into account their local circumstances. But it could also be argued, conversely, that the lack of any attempt even to make explicit what are considered to be the best practices artificially separates the definition of the European public interest by the Council from the identification of the means through which it may be realized—the very opposite of what a public policy based on the iterative redefinition of the objectives in the light of implementation should resemble. It is therefore crucial that the pooling of information through mutual information processes feeds into a debate concerning the further steps to be taken in the Union's legislative and policy agenda. Indeed, it is in the light of their contribution to the objectives pursued by the Union—rather than in the light of purely national preoccupations—that the measures adopted at Member State level should be evaluated: only if the information delivered by each Member State is examined in this light will such a process shape, in time, the attitudes of national actors, who instead of vetoing changes which would disrupt their expectations or acquired positions, could then become active participants in mutual learning aimed at the realization of the European public interest.

5 Conclusion

The lack of a clear, unambiguous understanding of the final shape of the AFSJ would constitute a disability under a classical, formalistic conception of policymaking: it would create an obstacle to the choice and sequencing of the measures to be adopted; and it would leave us ill-equipped when asked to define where mutual recognition can proceed, and where harmonization is instead required as a precondition. But the vague definition of the aims of the

AFSJ, and especially of their prioritization, can also be seen as an opportunity: perhaps counter-intuitively, it could encourage a mode of agenda setting more responsive to the actual needs of mutual cooperation between national authorities than to sudden events which, perceived as crises calling for urgent answers, may lead to an *ad hoc* and disorderly construction of the AFSJ. For this to happen, however, adequate mechanisms should be put in place. Monitoring of the situation of fundamental rights in the EU Member States is probably not a priority in this respect. Although such monitoring clearly would enhance mutual trust between the Member States, this is also an area where the standards are most uniform, and where the need for harmonization, therefore, may be weakest. If problems do occur—if, in other terms, diverging approaches to fundamental rights risk threatening mutual cooperation in AFSJ—courts generally may be counted upon to identify them: indeed, whether or not such problems exist depends largely on the attitude of courts, when they are confronted with allegations that mutual cooperation will result in a violation of fundamental rights.

But more is required, rather than less. First, the kind of monitoring which is required should go beyond fundamental rights strictly conceived; instead, it should ensure a screening of the developments within the Member States in all areas in which the European Union has launched policies, whether or not those developments relate to the implementation of a specific EU instrument, in order to identify the lacunae of existing policies and the need, therefore, to move further or even to change directions. Second, evaluation therefore should not be conceived primarily as a mode of supervision, or of monitoring compliance with pre-established commitments. To the extent that it includes a monitoring element, this should be limited to identifying instances where measures adopted in one Member State produce externalities, which may lead to the conclusion that some form of coordination, or possibly harmonization, is required. But the main purposes of such evaluation should be to promote *mutual learning* between the Member States and to ensure the *pooling of information on developments within the Member States needed to define the Union's agenda*. Any monitoring there is should be of a diagnostic nature: it should climb up the causality chain and identify which remedial measures might be suggested, and whether such measures should be adopted at the national or Union level.

The fulfilment of these objectives calls for as open and transparent a procedure as possible, as well as for the development of participatory mechanisms. Such procedures should ensure that the improved understanding gained from experiences conducted in other jurisdictions will not remain the privilege of certain high-level public servants involved in intergovernmental working groups, but instead will be diffused as widely as possible, in particular through umbrella non-governmental organizations or social actors established at European level. While it is often asserted that publicity can only operate at the expense of truthworthiness—since the national authorities may be tempted to

report only partially, highlighting their successes rather than their failures, if they know that the information they provide will be made public—any such tendency to misrepresent local conditions should be responded to by more transparency, rather than by more confidentiality: indeed, the preparation of shadow reports by non-governmental organizations may constitute a powerful incentive for states to provide as impartial and balanced a picture as possible of the problems they are facing, in order not to be accused of manipulating the facts. It is clear that in certain areas, confidentiality is required: where reports are presented about the control of the external borders, about counter-terrorism strategies or about the fight against organized crime, it is understandable that such reports should remain secret, since they may contain sensitive information that traffickers or terrorists, for example, might be able to use. But this will be true only in very exceptional situations, and in very limited fields. The fact that it is precisely in those areas that peer evaluations have been developed since a decade in the EU should therefore not be misconstrued: although, as we have seen, these evaluations are secret,³⁵ they are in this respect the exception, and should not constitute the norm for the future. In addition, while confidentiality may be a condition for a fully effective evaluation mechanism whose main objective is to ensure that Member States comply with certain requirements (such as, for instance, to control external borders, or to apply and implement at national level international undertakings in the fight against organized crime), since states in such a mechanism may have an interest in avoiding criticism, this justification is absent where the evaluation aims to promote mutual learning or to improve the relevance of EU policies and their ability to address the problems they seek to respond to.

While mutual learning and guidance of Union policies should be the primary objectives of evaluation mechanisms set up in the EU, such mechanisms at the same time should improve the accountability of national policy-makers and the quality of democratic deliberation. Openness and participation are conditions for mutual learning: the more different actors are involved, the less the national policy-makers directly in charge of any particular area will be able to afford to ignore the lessons from other jurisdictions and the easiest it will be for them to effectuate policy changes on the basis of those lessons, since the actors who could otherwise have vetoed or opposed such changes will themselves have been involved in this redefinition. Evaluation mechanisms thus conceived—based on the twin principles of publicity and participation—also have a deeply democratizing potential: they heighten the scrutiny to which national policy-makers are subjected, since the latter will have to explain both why their policies are failing where those developed elsewhere seem to work better, and why they are implementing certain policies despite their impact on other Member States. Such evaluation mechanisms also provide opposition political parties, civil society organizations, and the public at large with a broader range of options from which to choose and against which the policies

in force might be gauged. A virtuous circle may thus emerge, in which an improved evaluation of EU policies will have a democratizing effect at national level and, as a result, lead to improved national policies whether or not guided by a direct intervention from the EU.

For this to happen, constructing an evaluation mechanism at EU level along the lines of the proposals of the European Commission will not be sufficient. For the establishment of such a mechanism to produce the far-reaching impact we can hope for, certain background conditions should also be created, both at the European and at the national level. Member States must be convinced that it is in their interest to contribute to this evaluation process: that they can improve their policies by agreeing to discuss them with the other Member States, by asking how these policies contribute to the gradual shaping of the AFSJ, and by learning from the experiences of others. National parliaments should use this evaluation process as an opportunity to better monitor governments, who, as a result of this process, will have to provide justifications which they may not have had to provide previously; at the same time, the former will be led to redefine their understanding of what constitutes a valid justification, in the light of the impact on other Member States any particular choice made at national level may have. Civil society organizations should be active in this process, whose success will depend, to a large extent, on their vigilance and on their ability to feed into the evaluation the kind of grassroots knowledge they alone, in certain cases, may possess or may be willing to provide. Finally, specialized bodies, possessing a degree of expertise and independence ensuring that their opinions cannot be ignored or dismissed without justification by the government, could make an important contribution to such EU-wide evaluation processes conceived as tools for mutual learning.

Building an area of freedom, security, and justice is not to be conceived simply as the superimposition above national systems of governance, of another—European—layer, only marginally affecting the practices and ethos of the national authorities. On the contrary, because of the direct cooperation it requires between national authorities—law enforcement officers, national administrations, national judges—it could deeply transform those practices and ethos, and create among those concerned a sense of belonging to a new, broader, and more diverse community. It is our responsibility to ensure that this Europeanization of national practices in fields which, hitherto, were traditionally conceived of as belonging to the core of the sovereign powers of the state, results in greater accountability and in a richer democratic debate at both national and at European levels, rather than in the *déjà vu* impression of powerlessness of the masses in the face of European elites. Although, realistically, this darker scenario has greater chances of materializing in the next few years, this is by no means a necessity or the inevitable result of further powers being exercised at a level to which no *demos* corresponds. But we bear the burden of proving that there is an alternative.

Notes

1. Art. 6 (1) and (2) EU.
2. Art. 31 (1) (c) and (e) EU.
3. See Council Framework Decision of 13 June 2002 on the European arrest warrant and the surrender procedures between Member States (OJ L 190 of 18.7.2002, p. 1), 12th recital of the Preamble and Article 1, §3; Council Framework Decision 2003/577/JHA of 22 July 2003 on the execution in the European Union of orders freezing property or evidence (OJ L 196 of 2.8.2003, p. 45); Article 1, as well as, with regard to the *ne bis in idem* principle, the observance of which may constitute a ground for non-recognition or non-execution, Article 7 §1, c); Council Framework Decision 2005/214/JHA of 24 February 2005 on the application of the principle of mutual recognition to financial penalties (OJ L 76 of 22.3.2005, p. 16), 5th and 6th recitals of the Preamble as well as Articles 3 and 20 §3; the Council Framework Decision 2006/783/JHA of 6 October 2006 on the application of the principle of mutual recognition to confiscation orders (OJ L 328, 24.11.2006, p. 59), 3rd recital of the Preamble and Article 1, § 2.
4. Case 120/78, *Rewe-Zentral AG v. Bundesmonopolverwaltung für Branntwein* ('*Cassis de Dijon*'), [1979] ECR 649.
5. Joined Cases C-197/01 and C-385/01, *Gözütok and Brügge*, [2003] ECR I-1345.
6. Article 54 of the Convention implementing the Schengen Agreement (CISA) of 14 June 1985 on the gradual abolition of checks at the common borders, of 19 June 1990 (Schengen Convention), OJ L 239, 22.9.2000, p. 19.
7. The Court accepted that such mutual recognition would not be obligatory where it would jeopardize fundamental rights, such as the rights of victims of criminal offences; but it noted that, in the case at hand, the issue did not arise, since the *non bis in idem* principle 'does not preclude the victim or any other person harmed by the accused's conduct from bringing a civil action to seek compensation for the damage suffered' (par. 47).
8. For another example of this dialectic, see Council Framework Decision 2005/212/JHA of 24 February 2005 on confiscation of crime-related proceeds, instrumentalities, and property, OJ L 68 of 15.3.2005, p. 49, where the Preamble (10th recital) says that it is 'linked to a Danish draft Framework Decision on the mutual recognition within the European Union of decisions concerning the confiscation of proceeds from crime and asset-sharing, which is being submitted at the same time' (see now Council Framework Decision 2006/783/JHA of 6 October 2006 on the application of the principle of mutual recognition to confiscation orders, OJ L 328, 24.11.2006, p. 59, referring in turn to the Council Framework Decision on confiscation of crime-related proceeds, instrumentalities, and property).
9. The 1995 Data Protection Directive defines minimum safeguards for the protection of private life in the processing of personal data throughout the Union. Article 1(2) of the directive provides that 'Member States shall neither restrict nor prohibit the free flow of personal data between Member States for reasons connected with: the protection of the right to privacy with respect to the processing of personal data.
10. See Art. 3(2) of Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data, OJ L 281 of 23.11.1995, at 31.

11. C.E.T.S. no. 70; signed in The Hague on 28 May 1970 and in force since 26 July 1974. This instrument provides that each Contracting State shall be competent under certain conditions to enforce a sanction imposed in another Contracting State, which is enforceable in the latter State, upon the request of this State.
12. C.E.T.S. no. 73; signed in Strasbourg on 15 May 1972 and in force since 30 March 1978. Under the mechanism established by this Convention, any Contracting State may prosecute under its own criminal law any offence to which the law of another Contracting State is applicable, upon the request of the latter State.
13. See Article 6 of the European Convention on the International Validity of Criminal Judgments; and Article 11 of the European Convention on the Transfer of Proceedings in Criminal Matters.
14. See, for example, Article 11(2) of the Council Framework Decision of 13 June 2002 on the European arrest warrant and the surrender procedures between Member States (OJ L 190 of 18.7.2002, p. 1); Article 14(2) of the Council Framework Decision 2003/577/JHA of 22 July 2003 on the execution in the European Union of orders freezing property or evidence (OJ L 196 of 2.8.2003, p. 45); Article 16(2) of the Council Framework Decision 2005/214/JHA of 24 February 2005 on the application of the principle of mutual recognition to financial penalties (OJ L 76 of 22.3.2005, p. 16); Article 19(2) of the Council Framework Decision 2006/783/JHA of 6 October 2006 on the application of the principle of mutual recognition to confiscation orders (OJ L 328, 24.11.2006, p. 59).
15. Although principally based on the national provisions giving effect to the arrest warrant, as communicated to it by the Member States (as required under Article 34 (2) of the Framework Decision), the evaluation reports of the Commission (European Commission 2006c) also rely on the replies given to the European Judicial Network's questionnaire, which concerned the practical aspects of the arrest warrant prior to 1 September 2004, and by maintaining a bilateral dialogue with the designated national contact points.
16. Decision of the Executive Committee of 16 September 1998 setting up a Standing Committee on the evaluation and implementation of Schengen (SCH/ Com-ex (98) 26 def.), OJ L 239, 22.9.2000, p. 138. See Genson and van de Rijt (2006).
17. See Council of the EU, doc. 8286/1/03, of 6 May 2003 (Maintaining and increasing the efficiency of the Schengen Evaluation mechanism); and Council of the EU, doc. 15275/04, of 29 November 2004.
18. There are currently plans to further improve the evaluation mechanism (see the conclusions on the EU border management strategy adopted at the Justice and Home Affairs Council of 4-5 December 2006).
19. Common Manual (OJ C 313 16.12.2002, p. 97) and the Common Consular Instructions on visas for the diplomatic missions and consular posts (OJ C 313 16.12.2002, p. 1).
20. See Joint Action 97/827/JHA of 5 December 1997 establishing a mechanism for evaluating the application and implementation at national level of international undertakings in the fight against organized crime, OJ L 344, 15.12.1997, p. 7.
21. These have been mutual judicial assistance in criminal matters; the action of law enforcement authorities in the area of drug trafficking; the exchange of information between law enforcement authorities of the Member States and between the Member States and Europol; and the application of the European arrest warrant.

22. Council Decision of 28 November 2002 establishing a mechanism for evaluating the legal systems and their implementation at national level in the fight against terrorism (2002/996/JHA), OJ L 349, 24.12.2002, p. 1.
23. The theme of the first evaluation cycle was the exchange of information on terrorist activities.
24. Article 8(3) of Council Decision of 28 November 2002 establishing a mechanism for evaluating the legal systems and their implementation at national level in the fight against terrorism, cited above n. 22.
25. Article 61C of the Treaty on the Functioning of the European Union (as the EC Treaty will be renamed by the Treaty of Lisbon amending the Treaty on European Union and the Treaty establishing the European Community, signed at Lisbon, 13 December 2007 (OJ C 306, of 17.12.2007, p. 1)). Article III-260 of the Treaty establishing a Constitution for Europe (OJ C 310 of 16.12.2004, p. 1) had anticipated the introduction of such an evaluation mechanism.
26. Council of the EU, doc. 9778/2/05 REV 2, 10 June 2005; OJ C 198, 12.8.2005, p. 1.
27. This comprises six areas: external borders, visa policies and free movement of persons; citizenship and fundamental rights; coordination in the field of drugs; immigration and asylum; the establishment of an area of justice in civil and criminal matters; law enforcement cooperation and prevention of and fight against organized crime.
28. The Member States expressed their concern at what they consider the overambitious nature of the proposals of the Commission. In general, favour less frequent cycles of evaluation (every five years instead of twice every five years); and they advocated a focus, initially, on limited sectors, in order to 'test' the evaluation mechanism before extending it to all the policies covered by the communication. In addition, noting that the data-gathering techniques across the Union are not uniform, they questioned whether it was worth the effort reaching beyond the information already available in each Member State. See Council of the EU, 'Evaluation of Policies on Freedom, Security and Justice—Discussion Paper,' 8752/07 LIMITE, 23 April 2007; Finland's answers, 8752/07 ADD7 LIMITE, 29 May 2007; Sweden's answers, 8752/07 ADD9 LIMITE, 29 May 2007; Czech Republic's answers, 8752/07 ADD6 LIMITE, 29 May 2007; Answers from Republic of Slovenia, 8752/07 ADD1 LIMITE, 29 May 2007; Romania's answers, 8752/07 ADD3 LIMITE, 29 May 2007; Ireland's answers, 8752/07 ADD4, 29 May 2007; Poland's answers, 8752/07 ADD5 LIMITE, 29 May 2007; Austria's answers, 8752/07 ADD2 LIMITE, 29 May 2007; Denmark's answers, 8752/07 ADD8 LIMITE, 29 May 2007; Slovakia's answers, 8752/07 ADD10 LIMITE, 30 May 2007; Hungary's answers, 8752/07 ADD12 LIMITE, 31 May 2007; Belgium's answers, 8752/07 ADD11 LIMITE, 30 May 2007; Latvia's answers, 8752/07 ADD13 LIMITE, 31 May 2007; United Kingdom response, 8752/07 ADD14 LIMITE, 6 June 2007; Estonian answers, 8752/07 ADD16 LIMITE, 7 June 2007; Replies by the French delegation, 8752/07 ADD15 LIMITE, 6 June 2007; Greece's answers, 8752/07 ADD17 LIMITE, 8 June 2007; Portugal's answers, 8752/07 ADD18 LIMITE, 11 June 2007; Answers from Cyprus, 8752/07 ADD19 LIMITE, 11 June 2007; Bulgaria's reply, 8752/07 ADD21 LIMITE, 12 June 2007; Malta's reply, 8752/07 ADD20 LIMITE, 11 June 2007. I am grateful to Violeta Moreno Lax for collecting these answers.

29. The option of proposing the launch of an Open Method of Coordination was apparently considered in the course of the preparation of the communication on the evaluation of EU policies in the fields of freedom, justice, and security. This option was considered not to be politically feasible, however, since it was anticipated that the Member States would resist subjecting policies so closely linked to their national sovereignty to some form of peer review. In addition, the view was expressed that OMCs fit areas which are primarily inter-governmental in the absence of competences of the Union, whereas in the freedom, justice, and security fields covered by the evaluation proposed by the Commission, there exist EU policies and, increasingly, instruments implementing these policies (European Commission 2006e: 12).

30. European Commission 2000b: 2. The definition of evaluation is borrowed from the Glossary appended to the White Paper on Reform.

31. March and Olsen 1995: 227. In that sense, 'Democratic institutions (...) are both arranged to speed up and slow down learning from experience and adaptation' (March and Olsen 2001: 13).

32. One network currently monitors the implementation of Council Directive 2000/43/EC implementing the principle of equal treatment between persons irrespective of racial or ethnic origin (OJ L 180 of 19.7.2000, p. 22) and of Council Directive 2000/78/EC establishing a general framework for equal treatment in employment and occupation (OJ L 303 of 2.12.2000, p. 16); it is funded under the Community Action Programme to combat discrimination (2001 to 2006) (see Council Decision of 27 November 2000, OJ L 303 of 2.12.2000, p. 23). See: ec.europa.eu/employment_social/fundamental_rights/policy/aneval/legnet_en.htm. Another network of legal experts has been set up, since 1984, on the application of Community law on equal treatment between men and women. According to its web site, its tasks are to 'facilitate the Commission to monitor the implementation of the *acquis communautaire* in the Member states with regard to legislation and case law on equal treatment between women and men'; 'assist the Commission in the preparation of European Court of Justice cases, in particular infringement procedures'; and assist the Commission in developing 'new strategies and ideas' in this field: see www.ec.europa.eu/employment_social/equ_opp/rights/experts_en.html.

33. Existing academic networks, such as the European Criminal Law Academic Network (see ecjan-eu.org/) in the criminal law area or the Odysseus Network on asylum and immigration (see www.ulb.ac.be/assoc/odysseus/index2.html), could be used to that effect.

34. The European Commission initially had proposed the introduction in these areas of an Open Method of Coordination as an adjunct to the adoption of legislative measures under Articles 61–69 EC (see European Commission 2001c, 2001d). Since this met with scepticism, it then subsequently reverted to a more modest suggestion to enhance mutual information of national immigration and asylum policies between Member State policymakers through the creation of a mutual information procedure on planned national asylum and immigration measures, thus retreating from an OMC to information collecting and pooling (European Commission 2005e).

35. See above, Section 3.2.

12

Experimentalist Governance in EU External Relations: Enlargement and the European Neighbourhood Policy

Elsa Tuhets

1 Introduction

The external relations of the European Union (EU) are characterized by a patchwork of policies that are not always easy to coordinate. Some fields are communalized, like the Common Trade Policy and the negotiation of association agreements, where the EU Member States have agreed to transfer part of their sovereignty to the European level. Others still work intergovernmentally, for example, the Common Foreign and Security Policy (CFSP). Furthermore, as in the field of development, there is a single EU assistance policy, as well as (in theory) one policy for each Member State. The difficult tasks of how to coordinate such policies and thus to add coherence to European external relations are increasingly entrusted to the European Council (coordination between the Member States) and the European Commission (coordination between the Community level and the Member States). However, the literature on EU external relations overlooks the fact that these two institutions, especially the Commission, have recently helped to shape and implement a new method in European foreign policy. This method displays many similarities to the experimentalist forms of governance, developed during the 1990s in European internal policies like employment and social protection. I argue in this chapter that this link between internal and external policies was made possible through the process of EU enlargement, which extended internal policies to future Member States. This process of externalization of EU policies not only contributes to the Europeanization¹ of third countries interested in adopting EU norms and standards, but also gives opportunities for participation in European governance² through policy adaptation³ and the extension of internal (policy) networks abroad.⁴ This chapter focuses primarily