

Methodology



Evaluating European performance on the world stage for one particular year seems a reasonably straightforward exercise. The question, after all, is relatively simple: “Did Europeans do well or badly in 2010?” However, devising a methodology in order to make a rigorous and consistent judgment across issues and over time is a tricky enterprise that is fraught with unsatisfying trade-offs and inevitable simplifications. Before explaining the methodology used in this scorecard, we discuss some of the difficulties and dilemmas we faced while devising the methodology. This discussion is meant to offer some perspective on the choices we made and to ensure full transparency about the results.

Evaluating European foreign policy performance

Among the many difficulties involved with evaluating Europe’s performance in its external relations, two stand out: the problematic definition of success in foreign policy; and the rigidity of the time frame used.

What is a good European foreign policy?

The nature of international politics is such that “success” and “failure” are not as easily defined as they would be in other public-policy areas. In particular, there is no quantitative tool that can adequately capture performance in foreign policy as in economic policy or social policy (e.g. unemployment rate, crime rate, pollution levels, etc.). Diplomacy is more often about managing problems than fixing them, biding time, choosing the worst of two evils, finding an exit strategy, saving face, etc. States often pursue multiple objectives, and their order of priority is often unclear or disputed. This, of course, is even truer in the case of Europe, in which two member states might have different views on what exact mix of objectives met during the year constitutes success in one policy area, even when they agree on common objectives.

This difficulty is compounded by the heterogeneous nature of foreign policy. Europeans expect their authorities to solve the Israeli-Palestinian conflict, to prevent the proliferation of nuclear weapons, to turn Bosnia and Herzegovina into a functioning state, to protect ships from pirates in the Gulf of Aden, to stabilise the eastern neighbourhood, to defend European values at the UN and speak up for human rights, to convince other countries to fight climate change, to open foreign markets for exporters, to impose European norms and standards to importers, and so on. “Success” is defined very differently in each case: it can be a

matter of convincing other actors in a negotiation, building diplomatic coalitions, delivering humanitarian aid on the ground, imposing peace on a region torn by civil unrest, building a state, spreading global norms, etc. Moreover, Europe has very different abilities in each of them, not unlike the way that a student has different abilities in various subjects (e.g. mathematics, languages, physical education, etc.). This makes a unified grading system problematic by creating a dilemma between respecting the specificity of each “subject” on the one hand and ensuring that evaluations are comparable across the scorecard on the other.

Grading the rate of success of Europeans (the “outcome” score) relies on a comparison between the European objectives and the outcome for 2010. But the problem mentioned above resurfaces: who speaks for Europe? There is rarely a single entity to define what the European interest is – what priorities and trade-offs are desirable when conflicting objectives exist. Even where there is broad agreement on a policy, official texts will rarely present the real extent of European objectives, or will do it in vague, consensual terms. Therefore, simply comparing stated objectives with results would have led to an incomplete assessment of performance. It was generally necessary for us to go further and spell out explicitly what the European objectives were in one particular domain in order to compare them to results – a difficult and eminently political exercise.

What’s more, the causal link between one specific set of European policies on the one hand and results on the other is problematic. European objectives can sometimes be met regardless of the European policy put in place to achieve them. For example, independent factors might have modified the context in which actors operate (e.g. forest fires in Russia, rather than EU influence, led to a different attitude of Moscow towards climate change), or other states might have helped to attain the objectives sought by Europeans (e.g. the United States in getting China to support sanctions against Iran). But the opposite can also be true: failure can happen even with the optimal policies in place (e.g. the US Congress decision to abandon cap-and-trade legislation in spite of best efforts by Europeans to convince them otherwise).

This problem of causal disjuncture between policy and result led us to make two choices for the scorecard. First, we do not try to sort out the reasons for European “success”, let alone try to offer a co-efficient of European agency or credit. While we always specify other factors that contributed to a positive outcome, we deem Europeans to be successful if their objectives were met. In other words, they are not penalised for having been helped by others. This is why we use the word “outcome” rather than “results” or “impact” which imply a direct causality.

Second, we clearly separate policy from results. The grade for each component reflects an equal balance between input (graded out of 10) and outcome (graded out of 10) and output (graded out of 10), so that the reader can better appreciate the problematic correlation between the two. (The policy grade, or input, is divided into two scores, each graded out of 5: “unity” and “resources”.) Very good policies and best efforts can meet outright failure (e.g. the failure to get the US Congress to move on climate change). However, the opposite situation rarely occurs: luck, it turns out, is not so prevalent in international affairs.

Still, giving as much weight to policy as to results is a delicate choice that has several implications. It means that Europeans can get a score of 8, 9 or even 10/20 by having a policy we consider optimal, but a score of 0/10 or 1/10 for “outcome”. In other words, Europeans get a reasonably good grade for simply having a coherent policy in place, even if this policy produces few results. The other implication is that similar grades can mean different things. For example, on visa liberalisation with Russia (component 15), Europeans got 4/5 for “unity” and 3/5 for “resources” but only 3/10 for “outcome” – a total of 10/20. This is the same score as for relations with the US on counter-terrorism and human rights (component 31), where Europeans got 3/5 for “unity” and 2/5 for “resources” but a significantly better score of 5/10 for “outcome”.

Beyond the question of merits and results lies the question of expectations. If the scorecard has to spell out what European objectives were, it also has to define the yardstick for success, in the absence of obvious or absolute reference points to assess the underlying level of difficulty – and hence the level of success – in each area. We relied on judgment, based in each case on an implicit alternative universe representing the optimal input and outcome, against which actual European performance was measured. But while it was based on extensive expertise, this approach was necessarily subjective. This is particularly the case because, while it had to be realistic, it also had to avoid either lowering ambitions excessively or demanding impossible results. As noted in the Preface, this is where the political and sometimes even subjective nature of the scorecard is greatest.

It should also be noted that the relative nature of our judgment and the question of expectations contain an even more political question, that of European leverage – and, this time, the difficulty concerns both the policy score (i.e. “unity” and “resources”) and the results score (i.e. “outcome”). We evaluated performance in the context of 2010, and tried to be politically realistic about European possibilities, about what resources could be mobilised in support of a particular policy. But some observers might object that with some extra will or leadership

by the main actors, additional resources could have been mustered to increase European leverage, to the point of completely reconfiguring the political context of a particular issue. For example, on the Israeli-Palestinian conflict, some argue that Europe should take much more drastic and aggressive measures to reach its objectives. For example, it could unilaterally recognise a Palestinian state at the United Nations and bilaterally, or cease its Association Agreement with Israel and impose other trade sanctions. Admitting such proposals as realistic would change the score for “resources” (which, compared to this standard, would become dismal for 2010), and might potentially have changed the “outcome” grade as well. Here again, we had to make judgment calls about the adequacy of resources in the current European foreign-policy debate as we see it. It remains, however, a political judgment.

When does the clock stop?

A second set of problems has to do with the time frame of the scorecard. Evaluating foreign-policy performance is difficult enough, but it becomes even more difficult when you only consider events that took place during one calendar year. It is well known that some past policies that have yielded remarkable results in the short term proved less effective, and sometimes even disastrous, in the long term – for example, western support for the mujahideen in Afghanistan in the 1980s. The cost of some policy decisions has gradually increased over time – for example, the admission of Cyprus as an EU member state in the absence of resolution of the Northern Cyprus problem. Since the scorecard is an annual exercise, this will inevitably become an issue, especially after policies and actions we now vaunt prove less compelling in a few years, and vice versa. To some extent, however, this is the same problem we face in evaluating success not in absolute terms but as a function of possibilities and difficulty. We do not pass definitive historical judgment but rather a contextualised judgment within the bounds of the year 2010.

However, even that caveat does not solve the second dilemma: the possible bias in favour of short-term, tangible results that could be observed during the year 2010, to the detriment of more profound and meaningful, if less spectacular, policies and outcomes. For example, visa conditionality in the Balkans is exerting a continuing positive pressure and having good results, although these results are not evident on the larger, more visible political scene. The problem is that the scorecard tends to register movement, and while a European programme that is already in place can be mentioned in the text, it will often come second to the sometimes ephemeral

political battles that unfolded during the year. Thus, a limited but very visible political initiative towards a candidate country might eclipse the more important fact that the whole power relationship between Europe and this country is overdetermined by this candidacy. This bias is especially important when it comes to common security and foreign policy, since many aspects of the foreign relations of the EU take the form of long-term aid, development and rule of law programmes rather than short-term political initiatives. The scorecard tries to strike a balance between recognising the specificity, assets and successes of Europe as a different, new type of international power on the one hand, and considering Europe as a traditional great power, in the league of the US, China or Russia, on the other hand – a role it cannot escape in today's world.

This dilemma explains why, even though we insist on tangible results for 2010 and hold Europe to demanding standards of efficiency, we still give credit to and make room for patient background work and positions of principles, even if they seemed to have had no impact in 2010. After all, it was easy to criticise Europe for its failure to persuade the US to close Guantánamo prison until President Obama finally ordered its closure in 2009. It would be inaccurate to claim that the constant political and moral pressure that Europeans exercised played no role, and yet impossible to point out exactly what role they played in Obama's decision. Similarly, Europe's ongoing support of the development of the Palestinian Authority as a more effective and less corrupt administration is the type of behind-the-scenes work that is not always visible but could be hugely important in the future.

This question of time frame leads to the larger question of "good" foreign policies. We cannot assess whether policies are "good" – only whether Europeans are united around them, whether they devote resources to them, and whether (or to what extent) they reach their various objectives. In a sense, therefore, our judgment remains technical. For example, we find Europe's performance on Iran in 2010 to be better than on many other issues, but if Tehran suddenly acquires and uses a nuclear weapon in 2011, critics will point out that Europe's policy was not forceful enough and that the good grades we gave now look overblown. Similarly, if a revolution leads to the overthrow of the mullahs, critics will point out the immorality of European foreign policy that focused on the nuclear programme and reinforced the hardliners, while a more conciliatory position might have hastened the downfall of the regime.

This problem of normative judgment leads to a more general question: how much shall we take into account things Europe is not doing? For example, should

Europe get a bad grade because it was not present (in terms of either words or actions) in the China-Japan dispute of September 2010 about the Senkaku/Diaoyutai islands, where the future of world peace might be at stake? As discussed earlier, we have tried to strike a balance in the scorecard. On the one hand, we have graded existing policies and taken into account the specificity of EU foreign policy and what Europe actually is (i.e. long-term programmes and a certain vision of what the international system should be). On the other hand, we have graded according to “great power” norms, emphasising what Europe ultimately should be (e.g. an assertive power playing the multi-polar game).

The points above illustrate the difficulties and dilemmas involved in devising a methodology that can withstand criticism. This is why we call this project a scorecard rather than an index. Indices use hard quantitative data (e.g. UNDP’s Human Development Index; Brookings’ Iraq Index) or scores given by observers to qualitative data (e.g. Freedom House’s Freedom in the World or Freedom in the Press indices; Transparency International’s Corruption Perceptions Index), or a mixture of both (Institute for Economics and Peace’s Global Peace Index; Legatum Institute’s Prosperity Index). A scorecard, on the other hand, is transparent about the subjective nature of judgment and the heterogeneity of the material it grades, and is therefore a better tool for appraising foreign-policy performance. After all, the grades one gets in school are a function of the particular teacher doing the grading and are based on different criteria for each subject. However, this neither prevents the scorecard from being significant nor means that grades are purely arbitrary, especially when overall results are based on an average of a large number of exercises and as consistent a scale across the board and over time as is feasible.

Explanation of methodology

The scorecard was developed in three phases. In the first phase (during the summer and autumn of 2010), experts for each of the six “issues” drew up the list of “sub-issues” and “components” – the discrete elements that the scorecard actually evaluates for 2010. This choice, obviously, was fundamental as it determined what we were assessing within each of the six “issues” and was therefore the subject of intense discussion. The experts also provided preliminary assessments of European performance (for the period running from January to September) in each “component”, based on their own knowledge and a range of interviews with officials and specialists. In particular, they identified European objectives – a key precondition for evaluating performance. The experts devised questions for member states in order to better understand the dynamics of each component.

In the second phase (from November to December 2010), questionnaires on about 30 of the “components” on which the experts felt they needed additional information were sent to researchers in each of the 27 member states, who collected information from officials in their country and completed the questionnaires. This provided a much more granular image of European external relations on critical issues. In the third phase (January 2011), experts wrote the final assessments and the introductions for each issue. It was at this point that scores for each component were given. The scores and the assessments were then discussed with the scorecard team and shared with other experts and officials.

Criteria

The scorecard uses three criteria to assess European foreign-policy performance: “unity” (“Were Europeans united?”), “resources” (“Did they try hard?”), and “outcome” (“Did they get what they wanted?”). The first two evaluate the intrinsic qualities of European policies and are graded out of 5; the third criterion evaluates whether these policies succeeded or failed, and is graded out of 10. The overall numerical score out of 20, which was converted into an alphabetical grade, therefore reflects an equal balance between input and outcome.

In some cases, the scores for each of these three criteria are based on an average of several different elements of a “component”. For example, component 62, which evaluates European performance on Somalia, includes three disparate elements: the Atalanta naval mission; the training of Somali military personnel in Uganda; and financial support to the African Union peacekeeping mission AMISOM. Similarly, component 24, which evaluates relations with Russia on Afghanistan and Central Asia, has three elements: Afghanistan, Kyrgyzstan and security in Central Asia in general.

Unity

The key question on “unity” is: Do Europeans (that is, member states and EU institutions) agree on specific and substantial objectives or do they have a variety of different policies, with some adopting initiatives and taking stances that contradict the common policy?

Scores were awarded on the following basis:

- 5/5 = Perfect unity among member states and/or EU institutions – all agree on many objectives and push in the same direction(s). The best possible situation.
- 4/5 = A large degree of unity – member states and/or EU institutions agree on most objectives and positions but not all of them. Still a very satisfying situation.
- 3/5 = Partial unity, but member states and/or EU institutions have significant differences of approach and agreement exists on some objectives only. An acceptable situation.
- 2/5 = Strong differences in approach among member states and/or EU institutions – some take initiatives that contradict majority positions. An unsatisfying situation.
- 1/5 = A basic lack of unity among member states and/or EU institutions – there is no common agenda beyond a few common aspirations and conflicting positions dominate. A dysfunctional situation.
- 0/5 = Member states and/or EU institutions have opposite goals. In this situation, it is impossible to give a grade on resources and impact.

Some remarks:

- What is evaluated is not background harmony on a general issue such as Russia, but rather how united member states and EU institutions were on specific policy issues, events, initiatives or reactions in 2010. The context is not taken into account: unity is assessed in absolute terms, whatever the underlying level of difficulty. As a result, what could be called costly cooperation (i.e. cooperation attained in spite of deep underlying divisions) gets the same score as easy cooperation (i.e. cooperation attained because of already converging views).
- Process is not taken into account either: perfect or near-perfect unity on a range of objectives attained after stormy and protracted debates, and even disputes among member states and/or EU institutions, still justifies scores of 4/5 or 5/5 if the resulting policy line is observed by all, if all Europeans refrain from contradicting it in their external relations. Put differently, it means that misgivings, doubts, hesitations and silent disagreement among member states do not count. Only conflicting action is what is taken into account to evaluate and grade “unity”.

- Unity does not necessitate the existence of a common legal text or political declaration. Rather, the question is whether countries and institutions pushed in the same direction or not. If some abstained without hampering common action or making a difference, unity is still considered to be fully realised.
- Unity does not necessitate centralisation around Brussels. In other words, the scorecard does not have a normative bias towards a federal foreign policy, but it does have one towards a common and co-ordinated foreign policy.
- Unity is not an uncontroversial criterion of an effective European foreign policy. There is a case to be made that a lack of unity can either have no meaningful impact on results or even, in some rare cases, prove beneficial to Europeans. For example, while some argue that European division on the recognition of Kosovo limits its law enforcement actions in the Serbian-majority northern region and makes the EU less credible vis-à-vis Americans, others argue that the impact of European division is negligible or even positive (for example, because it means the EULEX mission is less intrusive and, therefore, improved relations with Serbia). Similarly, there are situations where European unity in multilateral forums (for example, in the form of a rigid and limited mandate) is an impediment to finding solutions and furthering European goals.

Resources

The key question on “resources” is: Did Europeans (that is, member states and EU institutions) devote adequate resources (in terms of political capital and tangible resources such as money, loans, troops, training personnel and the like) to back up their objectives in 2010? In other words, was their policy substantial?

Scores were awarded on the following basis:

- 5/5 = Member states and/or EU institutions devoted the largest possible resources imaginable in the real world (i.e. in the political, diplomatic, economic and budgetary context of 2010, not in absolute terms). They undertook bold initiatives, with the adequate expenditure of political, economic or military capital.

- 4/5 = Member states and/or EU institutions put serious resources behind the European position, but they were not quite as large or as bold as they could have been.
- 3/5 = Member states and/or EU institutions devoted only limited resources, with a negative impact on their ability to meet all the objectives.
- 2/5 = Member states and/or EU institutions devoted insufficient resources, leading to a clear gap between objectives and resources, which made it impossible for them to meet their objectives.
- 1/5 = Member states and/or EU institutions devoted few resources, resulting in a yawning gap between ends and means. If there was unity on objectives, then it was typically a soft consensus or was based on wishful thinking.
- 0/5 = Member states and EU institutions put no resources behind European positions.

Some remarks:

- Europeans can be only superficially united and agree on a purely declaratory policy. They can paper over the absence of meaningful unity by making lofty common declarations that are not backed by concrete action. They can, in a sense, “conspire” to hide their actual disunity behind joint declarations. Or, more frequently, they can reach a soft consensus on a course of action (or generally cosmetic action or non-action) which will result in a policy that cannot possibly make any difference in the real world. This is why this second criterion is added to the first. The “resources” criterion measures how substantial and ambitious European actions are – in other words, whether the policy is serious, whether it is backed up by resources and can make a difference or not, and how bold it is.
- Unlike the “unity” score, the “resources” score is assessed not in absolute terms but as a function of objectives and possibilities. It measures the gap between ends and means at a specific moment in time when material resources are not in infinite supply and when decision-makers have to make trade-offs between competing priorities. For each component, experts asked what other resources Europeans could have realistically devoted in order to reach their objectives. The score was determined by the gap between the reality of 2010 and the answer to this question.

- Therefore, this grade involves an eminently political judgment on what resources could realistically be mustered to support European objectives, whether they were adequate to meet them, but also, more profoundly, on how ambitious Europeans should have been. The remark made above about leverage is relevant here. If one thinks that Europeans ought to raise their game and adopt much more ambitious objectives on human rights in Russia and China, on stabilisation in Afghanistan or on visa reciprocity with the US, and mobilise additional resources to build extra leverage on these issues, one would award a lower score for “resources”. But in the scorecard we chose to base scores on objectives that are at the centre of gravity of the European consensus.

Outcome

The key question is: To what extent have European objectives been met in 2010, regardless of whether Europeans (that is, member states and EU institutions) were responsible for that outcome?

Scores were awarded on the following basis:

10/10 = All objectives have been met. There is a clear sense of success on this component (even in the case where Europeans cannot be credited for the entirety of that success).

9/10 =

8/10 = Most objectives have been met.

7/10 =

6/10 =

5/10 = Some objectives have been met. Disappointing results for Europe.

4/10 =

3/10 = No important objectives have been met. There were major setbacks for Europeans, and a sense of failure dominates.

2/10 =

1/10 =

0/10 = No objectives have been met. The outcome is the opposite of Europeans' aims, or the situation has deteriorated. A sense of uselessness/or even catastrophe predominates.

Some remarks (see also the section above on difficulties in evaluating European foreign-policy performance):

- While “outcome” assesses results, it does not attempt to measure success per se but rather success as a function of difficulty and possibilities, or performance given the underlying difficulty of the issues, or progress in meeting the objectives in the year considered. For example, it would be unfair and unrealistic to expect from Europeans that they single-handedly solve the Israeli-Palestinian conflict or stop Iran from enriching uranium. However, they can be expected to meet other partial objectives or make progress towards reaching them. For example, they can contribute to stabilising the Middle East or avoiding a sudden war, keeping the international community united, ensuring that the UN process is respected, or enforcing anti-proliferation norms.
- This criterion does not measure the European impact or Europe’s results, but the general outcome of the issue under consideration in the light of the initial European objectives. Many factors apart from European policies might have contributed to the 2010 outcome, including luck or a lack of it. While the scorecard always tries to indicate which other factors have played a role in a positive or negative outcome, it does not assess the outcome differently based on the perceived degree of European agency. In other words, in the case of a disappointing outcome, Europeans do not get a better grade because of adverse conditions, and in the case of a fortunate outcome, they are not penalised for having been helped by circumstances. Measuring the impact of European foreign policies would be a much more complex and hazardous exercise.
- European objectives or their degree of priority can sometimes change during a given year, which renders assessment difficult. For example, in 2009 and early 2010, Europeans wanted to convince Americans to shut down the Office of the High Representative in Bosnia and Herzegovina. While it remains an important goal for many of them, the events of 2010 led them to pursue this objective less forcefully.
- Defining the “outcome” criterion as “success as a function of difficulty and possibilities” leaves quite some room for divergent evaluations, as there is even less of a fixed yardstick than for “unity” and “resources”. Rather, the yardstick is redefined for each component in its proper context every year, in view of the European objectives during that year. This is where the political or even subjective nature of the exercise is most evident.

- However, judgment on outcome is not entirely relative or contextual. For each component, a balance had to be found between the relative or contextual scale (i.e. what objectives were met given the circumstances of 2010) and what could be called the absolute or ideal scale. Component 29 provides a good example of this: EU negotiators probably obtained the best possible deal they could from their American counterparts in the Open Skies negotiations on liberalising transatlantic air transportation, given their starting point. However, there remains a gross imbalance in market access in favour of the US, which is largely explainable by the inheritance of past bilateral deals with member states. In this case, Europeans got a good grade for their performance, but not the best possible one, since the overall result is still unsatisfying for Europe.

Numerical scores and alphabetical grades

Scores for “unity”, “resources” and “outcome” were added and converted into grades in the following way:

20/20	A+	Outstanding
19/20	A+	
18/20	A	Excellent
17/20	A-	
16/20	A-	Very good
15/20	B+	
14/20	B+	Good
13/20	B	
12/20	B-	Satisfactory
11/20	B-	
10/20	C+	Sufficient
9/20	C+	
8/20	C	Insufficient
7/20	C-	
6/20	C-	Strongly insufficient
5/20	D+	
4/20	D+	Poor
3/20	D	
2/20	D-	Very poor
1/20	D-	
0/20	F	Failure

Grades for issues and sub-issues

As indicated above, “components” are gathered in groups called “sub-issues”. The grade for a sub-issue simply results from the average of the grades for its components. Similarly, the grade for an issue such as crisis management or Relations with China simply results from the average of the grades for its sub-issues. This, of course, raises the question of the proper weight to grant to each component within a sub-issue, and to each sub-issue within an issue. For example, should the grade for China depend equally on the three sub-issues (Trade liberalisation and overall relationship; Human rights and governance; Cooperation with China on regional and global issues), or should one of them be granted more weight? Rather than engaging in a delicate exercise of weighting (for example, by giving co-efficients of importance to various components), we decided to build into the list a rough equality among components within a sub-issue and among sub-issues within an “issue”. It could be argued that some components and sub-issues have not been given their proper weight. However, such a judgment would be no less political than the grade given to that component.

COMPONENTS BY ISSUE	Unity (out of 5)	Resources (out of 5)	Outcome (out of 10)	Total (out of 20)	Score Grade
RELATIONS WITH CHINA				9.2	C+
Trade liberalisation and overall relationship				11.6	B-
1 Formats of the Europe-China dialogue	2	2	5	9	C+
2 Protection of European intellectual property rights in China	4	3	5	12	B-
3 Reciprocity in access to public procurement in Europe and China	4	2	3	9	C+
4 Trade and investment disputes with China	3	3	6	12	B-
5 Agreement with China on standards and norms, consumer protection	5	4	7	16	A-
Human rights and governance				5.7	C-
6 Rule of law and human rights in China	2	2	1	5	D+
7 Relations with China on the Dalai Lama and Tibet	2	1	2	5	D+
8 General openness of China on civil society exchanges	2	3	2	7	C-
Cooperation on regional and global issues				10.4	C+
9 Relations with China on Iran and proliferation	5	4	6	15	B+
10 Relations with China on Africa	3	3	4	10	C+
11 Relations with China on reforming global governance	3	2	2	7	C-
12 Relations with China on currency exchange rates	2	2	3	7	C-
13 Relations with China on climate change	4	4	5	13	B
RELATIONS WITH RUSSIA				9.5	C+
Trade liberalisation and overall relationship				11	B-
14 Trade liberalisation with Russia	4	3	5	12	B-
15 Visa liberalisation with Russia	4	3	3	10	C+
Human rights and governance				6.7	C-
16 Rule of law and human rights in Russia	4	2	2	8	C
17 Media freedom in Russia	3	2	1	6	C-
18 Stability and human rights in the North Caucasus	4	1	1	6	C-
European security issues				9.5	C+
19 Relations with Russia on the Eastern Partnership	3	2	3	8	C
20 Relations with Russia on protracted conflicts	3	3	4	10	C+
21 Relations with Russia on energy issues	4	2	3	9	C+
22 Diversification of gas supply routes to Europe	2	4	5	11	B-
Cooperation on regional and global issues				11	B-
23 Relations with Russia on Iran and proliferation	4	4	8	16	A-
24 Relations with Russia on Afghanistan and Central Asia	4	3	6	13	B
25 Relations with Russia on climate change	3	3	3	9	C+
26 Relations with Russia at the G20	2	2	2	6	C-

COMPONENTS BY ISSUE	Unity (out of 5)	Resources (out of 5)	Outcome (out of 10)	Total (out of 20)	Score Grade
RELATIONS WITH THE UNITED STATES				11.0	B-
Trade liberalisation and overall relationship				12.8	B
27 Reciprocity on visa procedures with the US	3	2	3	8	C
28 Relations with the US on terrorism, information sharing and data protection	5	5	8	18	A
29 Trade and investment disputes with the US	3	3	6	12	B-
30 Agreement with the US on standards and norms, consumer protection	4	3	6	13	B
Cooperation on European security issues				9.8	C+
31 Relations with the US on counter-terrorism and human rights	3	2	5	10	C+
32 Relations with the US on NATO and NATO reform	2	2	3	7	C-
33 Relations with the US on arms control and Russia	2	2	4	8	C
34 Relations with the US on the Balkans	3	4	7	14	B+
Cooperation on regional and global issues				10.6	B-
35 Relations with the US on the Middle East peace process	3	3	2	8	C
36 Relations with the US on Afghanistan	4	2	2	8	C
37 Relations with the US on Iran and proliferation	5	5	8	18	A
38 Relations with the US on climate change	5	4	2	11	B-
39 Relations with the US on global economic and financial reform	2	2	4	8	C
RELATIONS WITH WIDER EUROPE				9.5	C+
Western Balkans				13.3	B
40 Rule of law and human rights in the Western Balkans	3	4	6	13	B
41 Stabilisation of Kosovo	3	4	7	14	B+
42 Stabilisation of Bosnia Herzegovina	4	2	2	8	C
43 Visa liberalisation with the Western Balkans	4	5	9	18	A
Turkey				6.0	C-
44 Bilateral relations with Turkey	2	2	1	5	D+
45 Rule of law and human rights in Turkey	3	2	2	7	C-
46 Relations with Turkey on the Cyprus question	3	1	1	5	D+
47 Relations with Turkey on regional issues	2	3	2	7	C-
Eastern Neighbourhood				9.2	C+
48 Rule of law and human rights in the Eastern Neighbourhood	3	2	2	7	C-
49 Relations with the Eastern Neighbourhood on trade and energy	5	4	5	14	B+
50 Visa liberalisation with the Eastern Neighbourhood	3	2	5	10	C+
51 Resolution of the Transnistrian dispute	3	2	2	7	C-
52 Resolution of the Abkhazia and South Ossetia dispute	3	2	4	9	C+
53 Resolution of the Nagorno-Karabakh dispute	4	2	2	8	C

COMPONENTS BY ISSUE

	Unity (out of 5)	Resources (out of 5)	Outcome (out of 10)	Total (out of 20)	Score Grade
CRISIS MANAGEMENT				11.4	B-
Conflict Prevention and Mediation				10.7	B-
54 Crisis management in the Former Yugoslav Republic of Macedonia	2	3	6	11	B-
55 Crisis management in Sudan and Chad	4	3	4	11	B-
56 Crisis management in West Africa	3	3	4	10	C+
Humanitarian Action and Intervention				12.3	B-
57 Response to the earthquake in Haiti	4	4	8	16	A-
58 Response to flooding in Pakistan	3	3	5	11	B-
59 Response to the humanitarian crisis in Gaza	4	3	3	10	C+
Peacemaking and Peacekeeping Operations				11.3	B-
60 Stabilisation of the Georgian border	5	4	6	15	B+
61 Crisis management in Kyrgyzstan	4	1	1	6	C-
62 Crisis management in Somalia	4	4	5	13	B
State Building and Nation Building				11.2	B-
63 Stabilisation and state building in Afghanistan	2	4	3	9	C+
64 Stabilisation and state building in Iraq	5	2	4	11	B-
65 Stabilisation and state building in Bosnia and Herzegovina	4	4	5	13	B
66 Stabilisation and state building in Kosovo	3	4	5	12	B-
67 Stabilisation and state building in Congo	4	3	4	11	B-

COMPONENTS BY ISSUE		Unity (out of 5)	Resources (out of 5)	Outcome (out of 10)	Total (out of 20)	Score Grade
MULTILATERAL ISSUES					13.5	B+
European Impact in the Multilateral System					10.5	B-
68	European policy in the G20 and G8	2	3	5	10	C+
69	European policy on the reform of Bretton Woods institutions	3	3	4	10	C+
70	European policy on UN reform	4	2	3	9	C+
71	European policy on the financing of multilateral institutions	4	4	5	13	B
Human Rights and Humanitarian Issues					12.7	B
72	European policy in the Human Rights Council and UN General Assembly	3	3	4	10	C+
73	European policy on the ICC and ad hoc tribunals	4	4	7	15	B+
74	European policy in the international humanitarian system	2	4	7	13	B
Climate Change					15.0	B+
75	European policy on climate change in the multilateral context	4	4	7	15	B+
Non-Proliferation Regime					16.0	A-
76	European policy on Iran and proliferation in the multilateral context	5	5	7	17	A-
77	European policy on the NPT review conference	4	4	7	15	B+
Development and Trade					13.3	B
78	European policy on global health	4	3	6	13	B
79	European policy on the Millennium Development Goals	2	3	5	10	C+
80	European policy in the World Trade Organization	5	4	8	17	A-