Overseas Development Institute

Background Note

June 2012

Where do European Institutions rank on donor quality?

By Matthew Geddes

The impact of aid depends on a wide variety of factors such as its volume, the selection of the recipient countries, interventions and instruments, and the effectiveness of delivery.

The debate on aid impact is longstanding, with important contributions from Cassen, 1985, Mosley 1987, McGillivray, 2004 and Riddel, 2007. However, the tightening of budgets in the current financial crisis has led to a renewed focus on aid effectiveness, with the most recent iterations including three academic indices that rank the 'quality' of donors (Easterly and Williamson, 2011; Knack et al., 2011; Birdsall et al., 2010) as well as the Multilateral Aid Review (MAR, 2011) by the UK Department for International Development (DFID). These exercises are being used to assess donor comparative performance and foster international norms of good practice. The MAR is also being used to guide the allocation of DFID funds and identify areas of European Institution practice that DFID seeks to reform.

This paper investigates how to interpret, respond to and use the evidence they provide, focusing on the European Institutions, major donors themselves and, taken together, DFIDs largest multilateral partner (See Annex 1).

The paper has four sections.

1. The first section presents brief summaries of the scores that the four exercises award the European Institutions, finding that they are far from the bottom rankings and have particular strengths and weaknesses.
2. The second section reassesses this evidence and identifies issues that could make the evidence less robust.
3. The third section works through several use-case examples to see how the evidence that the indices present might best be applied.
4. Annexes provide detailed figures and tables on the four exercises.

Throughout this paper, European Institutions (EU Institutions) refers to the European Commission development budget, the European Development Fund (EDF) and the Humanitarian Aid and Civil Protection department of the European Commission (ECHO). EU refers to the EU Institutions and the member states. EC Budget refers to those funds managed by the European Commission but excluding those funded through the EDF. EC development budget refers to funds managed by the European Commission excluding both those funded by the EDF and humanitarian flows directed via ECHO.

The EC in current donor quality indices

All four exercises use indices to score the donors. To create an index, the authors define their understanding of effective / high quality aid and then select variables that represent the key components. These variables are typically proxies with, for example, the level of income per capita used to measure 'need'. Scores on variables are then combined through a system of weights to give each donor an aggregate score. This represents how well the donor's aid performs against the author's vision of what makes aid effective. The indices, and therefore the scores, vary across the exercises.

Easterly and Williamson (2011) measure donors' adherence to aid delivery best practices (measures of aid quality) as defined by aid agencies, outside aid monitors and academic literature. They claim no direct evidence of a link between these practices and aid impact. The five best practice dimensions are: transparency; overhead costs; specialisation/fragmentation; use of effective channels; and allocation to less corrupt/more democratic/more free recipients. Each dimension uses multiple indicators. The measure is of the absolute performance of donors, including trends over time and, for the most part, using equal weightings.

On this basis, and using 2008 data, the EU Institutions (titled European Commission) ranked as: average on specialisation; very good on selectivity; very poor on ineffective channels; below average for overhead; above average for transparency. Overall, they ranked 16th out of 42 donors, scoring 54%, i.e. just above average.

Knack et al. (2010) construct an aid quality index using 18 indicators drawn from the OECD Paris Declaration Monitoring Survey, OECD Development Assistance Committee (DAC) tables and aid data, and grouped into four dimensions of selectivity, alignment, harmonisation and specialisation with significant internal correlation of indicators. Several different weightings are provided to highlight the impact of changing weightings on overall rankings, with efforts made to control for country specific factors outside each donor's control.

Using 2008 data, the EU Institutions (combined with the European Investment Bank (EIB), together titled European Commission) ranked as: poor on selectivity; poor on alignment; good on harmonisation; good on specialisation. Overall, they ranked 23rd of 38 donors, scoring just below average.

Birdsall et al. (2011) investigate aid agency effectiveness (aid quality) for 23 donor countries and over 100 donor agencies, concentrating on measures over which aid agencies have control. A total of 31 indicators drawn from the literature are used as part of four dimensions of aid quality: maximising efficiency, fostering institutions, reducing burden, transparency and learning.

Using 2008 data, the EU institutions (combined with the European Investment Bank (EIB), together titled EC) are above average on all four components (see Table 1). They score better than the average for bilaterals on all four components, better than the UK for maximising efficiency and transparency and better than the average multilateral for fostering institutions and transparency. They scored worse than the World Bank International Development Association (IDA) – the only multilateral of comparable size – on all four components. As a group, multilateral agencies do better across the board.

Table 1: Birdsall et al. (2011) Z-scores (2008 data), higher numbers are better

	Maximising efficiency	Fostering Institutions	Reducing burden	Transparency	Average
EU Institutions	0.15	0.23	0.39	0.62	0.35
United Kingdom	0.08	0.51	0.54	0.23	0.34
Bilaterals	-0.16	-0.08	-0.14	0.00	-0.09
Multilaterals	0.43	0.23	0.39	-0.01	0.26
World Bank (IDA)	0.22	0.81	0.74	1.10	0.72

Source: www.cgdev.org/section/topics/aid_effectiveness/quoda and the author's own calculations.

The **Multilateral Aid Review (MAR, 2011)** uses a similar methodology to the three academic indices but adds three important components: the use of organisational theory to identify donor efficiency, an analysis of donors 'fit' with UK development objectives and the use of qualitative evidence, including from DFID staff and aid recipient countries. Overall, the EC Budget is rated as 'adequate' value for money for UK multilateral aid. The EDF and the ECHO are rated as 'very good'. The MAR scores multilaterals on an index with two main components: contribution to UK development objectives and organisational strengths comprising five and seven indicators respectively.

On its contribution to UK development objectives, the EC budget receives an overall score of 'weak' based on being rated 'satisfactory' in its criticality to UK and international development objectives; 'satisfactory' on climate change and environmental sustainability and fragile contexts; and 'weak' on gender equality, focus on poor countries and contribution to results. The overall score for organisational strengths is 'satisfactory' based on a rating of 'weak' for strategic and performance management and financial resources management, and 'satisfactory' for cost and value consciousness, transparency and accountability, and partnership behaviour.[1]

The EDF was rated 'strong' for both its contribution to UK development objectives and organisational strengths. Leaving aside those areas rated as satisfactory, indicators rated as strong are the EDF's criticality to UK and international development objectives and partnership behaviour. Weaknesses are gender equality and strategic and performance management. [2]

ECHO was rated 'strong' for contribution to UK development objectives and 'satisfactory' for organisational strengths. As with the EDF, leaving aside indicators rated as 'satisfactory', ECHO was rated as 'strong' for criticality to UK and international development objectives; focus on poor countries; and contribution to results. Indicators rated as 'weak' were strategic and performance management, climate change and environmental sustainability and gender equality.[3]

Table 2 summarises the strengths and weaknesses as defined by the four indices, focussing on the thematic dimensions that the indices develop rather than on the individual indicators they compose. See Annex 2 for the full results tables of all four indices.

The indices agree on several aspects, e.g. both MAR (2011) and Birdsall et al. (2011) find transparency to be a strength, and Easterly and Williamson (2010) and Knack et al. (2011) find 'specialisation' to be a weakness. However, in some cases the indices also disagree. Easterly and Williamson (2010) find 'selectivity' to be a strength, while Knack et al. (2011) find it to be a weakness.

Understanding the evidence

To use the evidence responsibly we need to investigate why it is possible for the indices to disagree. This section outlines several issues that might explain the differences and, in doing so, also suggests the need for deeper analysis where there is agreement.

Different data – are the indices comparing like with like?

The indices often use different data e.g. from different sources or years. These differences matter because donor scores can change significantly from one year to the next. There is also variation in the four exercises where the EDF and ECHO are included alongside the EC budget rather than being treated separately. Treating the EC as a single monolithic bloc is often unhelpful. The variations in scores between the EC's component parts are likely to be as wide as the variation between the EC and other donors. This is because the different components have very different roles and objectives and, therefore, approaches to aid delivery. Further confusion arises because ECHO and the EDF are managed by Europeaid while far less similar instruments (such as those for agriculture and enlargement that are managed by other European Directorates General) are rarely separated out.

Different indicators – why favour one set of indicators over another?

The four indices often use different methodologies to calculate indicators with similar names and concepts. Easterly and Williamson (2010) and Knack et al. (2011) agree that selectivity (conceptualised in a similar way) is important, but disagree on the best way to measure it (Box 1). The EC scores much better under Easterly and Williamson's definition.

Similarly, all four indices have their own beliefs about the key components of aid quality and these are reflected in their choices of dimensions. For example, not all of them include measures of transparency, and while greater transparency is widely thought to relate positively to aid impact, this link is not yet backed up by strong evidence. The EC tends to score well on transparency and, therefore, ranks higher on indices that

Table 2: Summary of scores in the academic indices and DFID's MAR

	Easterly and Williamson (2010) – EU Institutions	Knack et al. (2010) – EU Institutions	Birdsall et al. (2011) – EC Budget	MAR (2011) – EC Budget	Birdsall et al. (2011) – EDF	MAR (2011) - EDF
Strengths	Selectivity	Harmonisation	Transparency and learning, Reducing burden		Fostering institutions, Reducing burden	Partnership behaviour
Average	Overhead, Transparency	Alignment	Fostering institutions	Cost/value consciousness, Partnership, Transparency and accountability,	Transparency and learning	Focus on poor countries, Financial resource management, Transparency and accountability, Cost and value consciousness
Weaknesses	Specialisation, Ineffective channels	Selectivity, Specialisation	Maximising efficiency	Focus on poor countries, Strategic and performance management, Financial resource management	Maximising efficiency	Strategic and performance management,
Overall	Average	Average	Above average	Satisfactory	Above average	Strong

Note: For comparability, the MAR(2011) results shown here focus on the organisational aspects with the addition of the 'focus on poor countries' indicator.

Box 1: Selectivity

For Easterly and Williamson (2010) the concept of Selectivity refers to: '*aid delivery to the poorest countries while avoiding corrupt dictators.*' They state that '*aid is more effective at reducing poverty when it goes (1) to those countries in most need of it (the poorest countries), and (2) to countries with democratically accountable governments, and (3) less corrupt governments*.'

Knack et al. (2011) have a similar definition: '*Aid is widely believed to have greater development impact where it is needed most – that is, where there are large numbers of poor people – and where the policy and institutional environment is favourable to growth and development.*'

With such similar conceptualisations, both papers try to measure the same relationship. Therefore the significant difference in scores must be related to the measurement methodology. Easterly and Williamson (2010) say: '*we calculate the share of aid going to low-income countries, free countries (based on democracy scores), and less corrupt governments. We create an overall composite selectivity score where donors get positive weight on aiding poor countries and negative weight on supporting corrupt or un-free countries*.'

Knack et al. (2011) say: '*we regress the log of aid from donor i to recipient j on three variables for the recipient j: the log of its population, its GDP per capita (adjusted for purchasing power parity), and its overall score on the World Bank's annual "Country Policy and Institutional Assessment" (CPIA). The CPIA consists of 16 indicators "representing the different policy and institutional dimensions of an effective poverty reduction and growth strategy."*'

The indices pick different measures because there is no conclusive evidence as to whether the CPIA or an index of democracy is a better predictor of a positive relationship between aid spending and aid impact. Both the CPIA and the democracy index are themselves used as rough proxies for 'good policy' – a factor that, despite much investigation, remains heavily contested on whether 'good policy' can be an important determinant of aid impact, what 'good policies' are, and in which contexts those policies are valid.

include it as an indicator. The result of this uncertainty is that it is hard to decide how well the EC scores. There is no obvious choice between the pieces of conflicting evidence on which to base decisions.

An unclear link to aid impact – the indicators used are not always a good proxy for aid impact

It is hard to decide which set of indicators is a better measure, given the lack of robust evidence linking the indicators of aid quality to aid impact. This is summed up by Knack et al. (2011) who state that 'most indicators of donor performance are based on plausible but largely untested beliefs about best practices in aid management'. This uncertainty often makes the resulting use of indicators problematic.

- Many aspects of aid quality are not easily measured. For example, to have a positive impact, specialisation needs to take place according to each donor's comparative advantage. No data on this are available, so the indices assume that any specialisation that takes place follows each donor's comparative advantage automatically, but with little evidence to demonstrate that this is the case.
- Little account is taken of external factors. While alignment and harmonisation are often desirable concepts, for example, there is no practical guidance on what donors should do if the recipient government does not lead development efforts. Aligning to a government that is not interested in poverty reduction would not, presumably, lead to poverty reduction.
- Where there is evidence on aid effectiveness, it is often concerned with 'what not to do'. In the indices, however, the logic is often reversed to suggest that the opposite outcome is always better. For example, while high levels of technical cooperation are assumed to be inefficient, it seems likely that zero technical cooperation would also be inefficient but this is not accounted for in the calculations. The indices also struggle to identify the appropriate score where indicators are characterised by non-linear relationships.
- There has been little or no testing of whether the 'best practices' form a cohesive blueprint – do they work well together or are there trade-offs between the components? For example, lowering administrative costs may mean that there are not enough staff to deliver aid in the poorest countries where government capacity is weaker. As such, it is not clear that the different indicators are suitable for aggregation into a single index. All three academic indices recognise this and caution against the creation of overall rankings, preferring a separate ranking for each key concept with end users identifying the specific strengths and weaknesses of each donor.
- None of the rankings take into account changes in optimal aid quality behaviour that might take place in relation to the behaviour of other donors. If all donors stopped funding relatively better off countries or global public goods to concentrate on the very poorest countries, each donor would be ranked more highly but the global outcome would not be desirable.
- The indices focus on aid quality but rarely adjust for the volume of aid delivered. When a donor's score for each recipient country is aggregated, the score for donors like the EU Institutions with mandates to be present in the majority of aid recipient countries (and

therefore having many low volume country programmes) will be biased and may not represent the majority of aid they deliver.

- Little evidence is provided to demonstrate that indicators are aggregated in a way that reflects their relative importance for achieving aid effectiveness. Typically, the indices weight indicators equally for lack of alternative hypotheses but this is unlikely to reflect reality. For this reason, all three academic indices caution against creating a single aggregate index.
- The effectiveness of an aid delivery practice is specific to each country, sector, intervention and mandate i.e. a practice that is effective in one situation may not be so in another. For example, while using recipient government procurement systems is seen as best practice, using donor systems might be better when working in fragile states with less capacity or on humanitarian interventions that rely on speed of delivery. The three academic indices struggle to take this into account as it would be too complex to address this with formulae. As a result, they probably penalise the EU Institutions for working in fragile states where it is either more difficult or not appropriate to provide aid on the basis of a best practice designed for non-fragile countries. Similarly, the indicators may not work well for the less traditional instruments used by the EU Institutions such as blended finance.

One approach to address many of these issues is to incorporate other types of evidence that are far less affected by such challenges. Box 2 describes the MAR approach.

The wider picture

The challenges discussed in this paper make it difficult to say whether the EC's ranking in the indices is actually a good indicator of the impact the EC has on the ground. There are many examples of practices that lead to greater aid impact that are not captured – or that are even penalised – by the indices. Despite these issues, however, the indices contain valuable evidence. The challenge is to use them appropriately. Two examples are given below: a more general look at aid allocation and an issue that is topical for the EC – administrative costs.

Box 2: Stakeholder perspectives and other forms of evidence

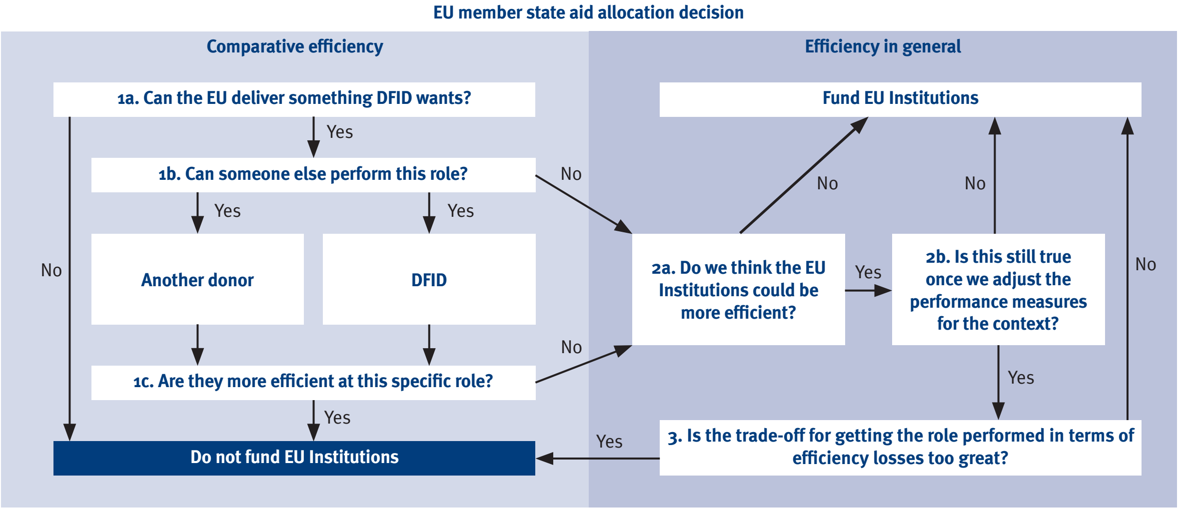
The best measure of aid quality would be the impact on the recipients. Unlike the three academic indices that use only indirect proxies for aid impact, the Multilateral Aid Review (MAR) supplements them with evidence from people and organisations that observe aid's impact directly. The MAR (p.108) reasons:

'*Most academic assessments of donor effectiveness are based on a set of quantitative indicators, such as disbursement delays, or adherence to Paris commitments. This approach has the advantage of being very clear and transparent. But it also has disadvantages. First, there are no data sets which cover all of the multilateral organisations… Second, some of the criteria that we were concerned about are not very amenable to this approach. For example, it would be extremely difficult to develop a set of measurable quantitative indicators which effectively capture the answer to the questions, "Does the multilateral organisation challenge and support partners to think about value for money?" or "Are its objectives sufficiently challenging?". Furthermore quantitative indicators do not easily allow us to draw on the knowledge that DFID and other parts of the UK Government have built up through a long history of continued engagement with these organisations.*'

Stakeholder perceptions are an established tool. Examples of approaches that use such perceptions include the Multilateral Performance Assessment Network (2011), the Humanitarian Response Index (2011), Debt Relief International (2004), Burall (2007) and Keystone Accountability (2011). Commonly perceived EC Budget advantages include: being 'joined up', its accountability, its wide range of instruments, its trade link, its use of partnership arrangements and its country presence. Negative perceptions are that the EC Budget is slow and over centralised. Few of these aspects are captured in the academic indices. Evidence from the Listening Project (2011) and Wathne (2009) also support this suggestion that the perspective of aid recipients on the key indicators for aid effectiveness differ from the indicators that make up the international consensus. DFID collated this information through country visits and stakeholder consultations for the MAR, reasoning that: '*One of the most important tests of the effectiveness of multilaterals is how country stakeholders view them. Government officials and civil society representatives who work on the ground with multilateral organisations have authentic judgments based on their day to day experiences and first hand observations of impact.*'

DFID also used evidence from existing multilateral evaluations such as DAC peer reviews, in-depth, programme and academic evaluations and results-based management reports, recognising the value in triangulation between these sources. DANIDA (2008) refers to this as a synthetic approach – one that is vital if the desired outcome is a 'bottom line' with which to answer wider questions about comparative performance, efficiency and allocation. However, it is unclear how systematically these other sources were integrated. Finally, DFID's separate bilateral aid review introduced 'results-offers' – a methodology based on assessments by donor staff of expected aid impact that might also work for the assessment of multilateral allocations.

Figure 1 - Policy influencing approaches



Aid allocation

Figure 1 outlines the potential role the indices can play for DFID's aid allocation question in the MAR, nominally: does money spent through the EU Institutions meet DFID's value for money, transparency and accountability requirements? But also: what development objectives can DFID best pursue through the EU Institutions rather than bilaterally or via another multilateral?

The academic indices can only help with question 2a in Figure 1. If they vastly increased the level of detail they could also help to answer question 1c and 2b. To make an allocation decision requires the additional information provided by the MAR for stages 1a and 3. For stage 3, the evidence for where DFID draws the line is entirely opaque but the line is clear: were the EU Institutions to fall below a set score, despite the lack of alternatives for some EU Institutions roles, DFID would judge the efficiency losses compared to spending the funds elsewhere to be too great.

Figure 1 highlights the context specificity required to answer practical questions. In their current form the indices cannot suggest whether DFID should outsource its budget support to the EU Institutions as the rating given is based on total EU Institutions aid, of which budget support operations form a minority. The aggregate score obscures whether the comparative strengths and weaknesses of the EU Institutions make it better at delivering budget support than DFID.

Another feature of Figure 1 that is often confused in the indices is that it is essential to separate criticisms of the value judgements that the index is designed to reflect from criticisms that the way an index is constructed doesn't embody those values (separating stages 1a and 1b from the rest). One example of a value judgement is how much weight the MAR gives to poor populations (an exponent of 0.2 in the needs-effectiveness index). The expectation is that DFID justifies this choice based on evidence but, in the end, this is a judgement. Unfortunately, the construction of the indices means that the values they embody in their construction are often either opaque or appear to be almost accidental.

An extreme example is the MAR's treatment of the United Nations Industrial Development Organization (UNIDO), which scores badly in the MAR because it is mandated to perform a role that DFID does not value. While many economists might disagree, if DFID does not value UNIDO's contribution, it should be clear how much this value judgement affects UNIDO's overall rating. However, UNIDO also scores badly because it doesn't meet DFID's criteria of working in fragile states and because its programmes are mostly technical assistance and, therefore, small in terms of aid volume. This analysis seems misguided, with the MAR failing to separate role from efficiency. UNIDO's support is not appropriate for fragile states, so the fact that it does not work in fragile states should have a positive impact on UNIDO's score, not negative. The fact that UNIDO does not count as 'critical' because its programmes are too small is not a sensible analysis: technical assistance programmes are always small compared to, say, infrastructure or budget support.

While the MAR attempts to judge donors with criteria applied flexibly and according to their role, the desire to compare donors that perform very different tasks can still produce perverse results where proxies for impact are used, such as working in fragile states and having large programmes. In the case of the EU Institutions, the MAR's flexibility to capture the types of benefits that the academics are unlikely to capture may well be important.

For example, the EC is mandated to have a global field presence, including in upper middle income countries, from which DFID benefits, and the EC can also mobilise large scale funding and unique instruments and bring specialist expertise alongside political leadership and greater legitimacy in conflict situations. How are these valued e.g. when compared to the additional costs they impose on the EU Institutions? These value judgements, whether intentional or unintentional, are often very obscure by virtue of being built into, and therefore hidden within, the twin black boxes of data collection and index mechanics.

Administrative costs

The EC Budget is often challenged on its administrative costs on the basis that it would represent greater value for money if these were lower. The general consensus across the four exercises is that the EC Budget has moderate to low administrative costs. Birdsall et al. (2011) find that of the 30 donors surveyed, the EU institutions (including the EIB) had the lowest unit administrative costs; Easterly and Williamson (2010) find that the EU institutions' overhead costs (which include administrative costs) are within the expected range; Knack et al. (2011) found that the administrative costs (indicator sp2) are approximately average and the MAR reports that EC Budget administrative costs are moderate.

Administrative cost data are not very comparable. The indices use different data, with ECHO, for example, either included or not, and the indicators they use are different, with some indices adjusting for donor size, hypothesising that there are some fixed costs that shouldn't be included in a measure of efficiency. The reporting of administrative costs is not standardised across institutions so it is unclear if the items under 'administrative costs' are the same and, therefore, comparable for each organisation. DFID, for example, is under political pressure to reduce recorded administration costs. Anecdotal evidence suggests that this is achieved in part by reclassifying administration functions into programme budgets, which reduces reported costs.

When donors outsource aid delivery to multilateral or bilateral partners, this transfers the bulk of the administration costs for aid delivery but does not transfer the associated volume figure that represents the denominator in the measure 'administration costs as a share of ODA'. DFID spent approximately 50% of its budget through partners in the year 2010/11, while the EU Institutions spent 14% of their budget through other donors (OECD, 2010), which significantly reduces DFID's administrative costs compared with those of the EC Budget.

Administrative cost measures pose problems for index construction. Even if administrative costs were standardised in their presentation, this would not necessarily produce straightforward comparisons as lower costs do not necessarily result in better value for money, and may often result in reduced aid impact. Spending (as in administrative costs as a share of ODA) is a poor proxy compared to the ideal proxy of administrative costs per unit of aid impact. It is a poor proxy because administrative costs and aid impact vary according to many factors, including the number of countries with programmes, aid instruments used, typical volume of support per intervention, legal requirements in both donor and recipient countries, the local environment (i.e. working in fragile states), and others. For example, fragile states require a high level of monitoring, supporting and evaluation for aid to be effective. This makes aid in fragile states more expensive to deliver than aid in more stable contexts but it might have more impact.

Similarly, the EC Budget is present in many more countries in the world than DFID (in 143 countries compared with DFID's 28 countries). Maintaining country offices is a significant cost but one that confers significant capacity to achieve greater impact. Bigsten et al. (2011) calculate that reducing the number of countries by 37% would reduce administration costs by 20%. As a result, the UK can respond more effectively through EC Budget field offices in countries facing famine in the Sahel where DFID itself does not have a presence.

The indices are probably not a good source of evidence on administrative costs. It is unclear whether they use comparable data, and there is insufficient evidence describing the relationship between administrative spending and aid impact. The indices do not factor in whether reduced administrative costs would affect the EC Budget's ability to use its comparative advantages e.g. a strong commitment to partnership, its coordination role and its focus on fragile states and humanitarian work or more basic tasks, such as producing research and policy advice and implementing projects.

If administrative costs were lower for one donor than another, and nothing else was different, then this would be an improvement. But this condition (as assumed by the indices) seems unlikely. Further evidence is needed to understand how administrative costs and the capacity they can confer relate to project outcomes (impact). Are the administrative costs high given the capacities (e.g. research and policy staff) of donors and the outcomes that they achieve? Evaluations of administrative costs exist and a better starting point might be for the indices to compare similar interventions across donors to isolate administrative efficiency from the impact of, for example, the type of intervention used.

Conclusion and proposals for further research

It is important to ask questions on the effectiveness of aid delivery, and the three academic indices and the MAR are a valuable source of evidence with which to answer them. They have already led to considerable changes in the way donors are compared and aid is allocated.

Their results suggest that the EU Institutions do some things very well and other things less well. However, the conflicting nature of the results they provide is not helpful for policy-makers, with the indices struggling, in part, because they draw upon a field where there is a general lack of robust evidence.

The highly complex problem of linking donor practices to aid impact is probably not a problem best suited to an index approach. On their own the indices are limited in what they can be used to say robustly, even more so when being used to produce a bottom line with which to allocate aid funding. Some problems could be eased by incorporating stakeholder perspectives, increasing the level complexity and the level of transparency. But the remaining weaknesses of the indices mean that they are best used as one source among many, synthesising the results from the indices with results from other types and sources of evidence. They stimulate some interesting questions, but leave many unanswered.

The practical challenge in their use is identifying where the indices live up to their design and where unintended consequences of their design obscure the trends that are being sought. Making best use of the evidence the indices provide involves checking their definition of a concept, understanding the nature of their links to aid impact, knowing what data are being used and making sure that the indicators are valid in the context being investigated. The evidence they provide will always require the user to make value judgements about the trade-offs between different aspects of donor behaviour and users will often have to consider factors beyond the indices model, such as political feasibility. The MAR incorporates several of these additions to transform an index into a useful tool for aid allocation, but it is still imperfect.

The international aid system needs devices that highlight where aid donors are displaying good practice or may have room for improvement. However, for a donor agency that has received a ranking, the indices alone are not advanced enough to provide the specific policy advice or recommendations necessary to stimulate a change in practice. The research agenda required for this will involve producing indices with supporting case study evidence at a disaggregated level to address specific issues. Perhaps, as happens in other fields, a better role for indices is to evaluate the system as a whole and identify different trends in, or groupings of, donors that might be taken up by other methods.

Written by Matthew Geddes, ODI Research Officer (m.geddes@odi.org.uk).

The author would like to thank Siân Herbert, Mikaela Gavas, Simon Maxwell and Heidi Tavakoli for their helpful comments. This paper is part of ODI's European Development Cooperation Strengthening Programme (EDCSP), funded by DFID. See http://international-development.eu

References and endnotes

References

Bigsten, A., Platteau, J., Tengstam, S. (2011) '*The Aid Effectiveness Agenda: The benefits of going ahead*'. Turin: SOGES. http://ec.europa.eu/europeaid/how/ensure-aid-effectiveness/documents/aid-effectiveness-benefits_en.pdf

Birdsall, N., Kharas, H., Mahgoub, A., Perakis, R. (2010) '*Quality of Official Development Assistance Assessment*'. Washington: Centre for Global Development. http://www.cgdev.org/files/1424481_file_CGD_QuODA_web.pdf

Burral, S., Mease, K., Pooja, M., Datta, A., Kamau, N. (2007) '*Assessing Key Stakeholder Perceptions of the Effectiveness of Multilateral Organisations*'. London: Overseas Development Institute. http://www.odi.org.uk/resources/docs/239.pdf

Casson et al. (1985) Does aid work? Oxford: Clarendon Press.

DANIDA (2008) '*Assessing Multilateral Organisation Effectiveness*'. Copenhagen: DANIDA. http://www.oecd.org/dataoecd/48/34/42211723.pdf

DFID (2011) '*Multilateral Aid Review*'. London: Department for International Development. http://www.dfid.gov.uk/Documents/publications1/mar/multilateral_aid_review.pdf

Debt Relief International (2004) '*The Effectiveness of Aid to Africa Since the HIPC Initiative: Issues, Evidence and Possible Areas for Action*'. London: Debt Relief International

Easterly, W. and Williamson, C. (2011) *Rhetoric vs Reality: The Best and Worst of Aid Agency Practices*. World Development, Volume 39, Issue 11, November 2011, Pages 1930–1949. http://www.sciencedirect.com/science/article/pii/S0305750X11002026 (accessed at time of writing from http://dri.fas.nyu.edu/docs/IO/19960/Rhetoricvsreality.pdf)

Geddes, M. (2012) '*Adjusting indices of donor quality*'. London: Overseas Development Institute. http://www.odi.org.uk/resources/docs/7694.pdf

HRI (2011) '*Humanitarian Response Index*'. Madrid: Dara. http://daraint.org/humanitarian-response-index/humanitarian-response-index-2010/information-sources/

http://www.hipc-cbp.org/files/en/open/Advocacy/DRIdocs/DFI_Aid_Effectiveness.pdf

Keystone Accountability (2011) '*NGO Partner Survey 2010*'. London: Keystone Accountability. http://www.keystoneaccountability.org/sites/default/files/Keystone%20partner%20survey%20Jan2011_0.pdf

Knack, S., Rogers, H., Eubank, N. (2011) *Aid Quality and Donor Rankings, World Development*, Volume 39, Issue 11, November 2011, Pages 1907-1917 http://www.sciencedirect.com/science/article/pii/S0305750X11002038 (accessed at time of writing from http://www-wds.worldbank.org/servlet/WDSContentServer/WDSP/IB/2010/05/04/000158349_20100504103946/Rendered/PDF/WPS5290.pdf

Listening Project (2011) '*Listening Project and Aid Effectiveness: Aid Recipient Perspectives on the Paris Declaration*'. Cambridge: CDA. http://www.cdainc.com/cdawww/pdf/other/perspectives_on_the_paris_declaration_Pdf_1_1.pdf

McGillivray (2004) '*Is Aid Effective?*', UNU-WIDER: Helsinki.

Mosley, P. (1987) Foreign aid: Its defense and reform. Lexington: Univ. Press of KY.

Multilateral Operational Performance Assessment Network (2011) '*MOPAN Survey 2011*'. Montreal: Universalia. http://www.mopanonline.org/ the 2008 Assessment of the EC can be found online http://www.diplomatie.gouv.fr/fr/IMG/pdf/Annual_MOPAN_survey_2008-3.pdf

OECD (2010) '*2010 DAC Report on Multilateral Aid*'. Paris: OECD. http://www.oecd.org/dataoecd/23/17/45828572.pdf

Riddell, R. (2007) Does foreign aid really work? New York: Oxford University Press.

Wathne, C. (2009) '*Multilateral Aid Organisations: Stakeholder Views on Effectiveness*'. London: Overseas Development Institute. http://www.odi.org.uk/resources/download/3772-stakeholder-effectiveness.pdf

Endnotes

1. http://www.dfid.gov.uk/Documents/publications1/mar/ecbudg.pdf
2. http://www.dfid.gov.uk/Documents/publications1/mar/edf.pdf
3. http://www.dfid.gov.uk/Documents/publications1/mar/echo.pdf
4. http://ec.europa.eu/development/icenter/repository/european_consensus_2005_en.pdf
5. http://ec.europa.eu/europeaid/what/development-policies/documents/agenda_for_change_en.pdf

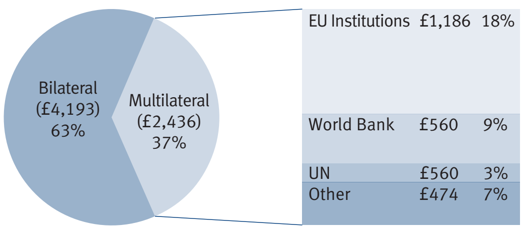
Annex 1: The EC and DFID

The European Institutions are a key channel for the UK's international development programme, responsible for channelling 49% of DFID's multilateral budget in the financial year 2009/10, representing 18% of DFID's total expenditure (see Figure 1).

These funds, alongside those from the other member states, create two sources of funding for EC external assistance: the European Commission budget (EC budget) and the European Development Fund (EDF) and represent 35% of global multilateral inflows. Together they committed €11.1 billion in 2010 (of which €9.9 billion qualified as ODA) making the EC the world's largest multilateral donor and the fourth largest donor behind the US (See Figure 2). Figure 3 splits the funds by their management within the EC.

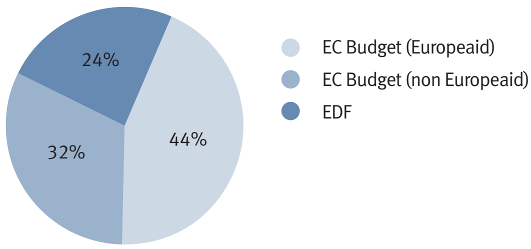
The EU institutions carry out the mandate agreed by their contributors – the member states – and their development policy is governed by the European Consensus on Development.[4] The next two years will be crucial in reforming EC development policy, given negotiations over the European Commission's new proposed development strategy 'An Agenda for Change'[5], the Multi-Annual Financial Framework (MFF) 2014-2020 and the EDF replenishment.

Figure 1: DFID's financial contribution (millions) to the EC in FY 2009/10



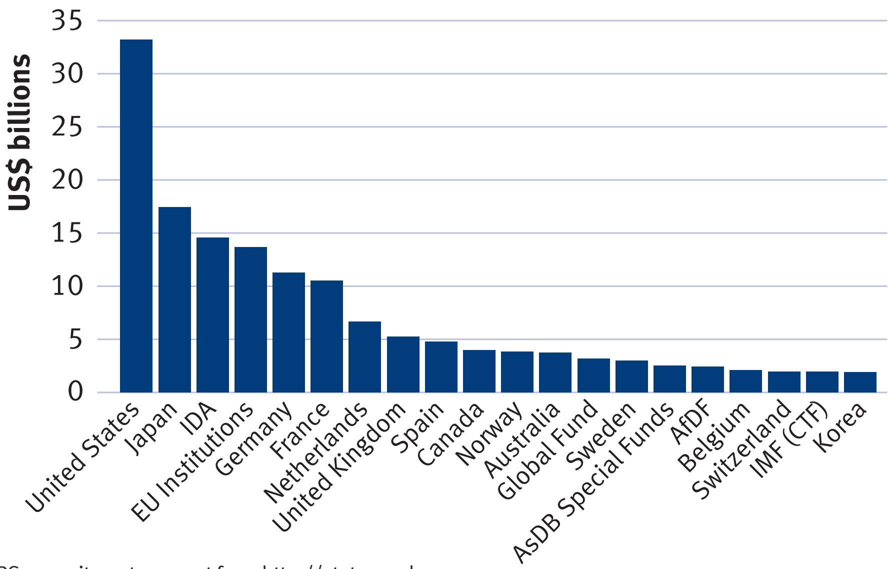
Source: www.dfid.gov.uk/About-us/How-we-measure-progress/Aid-Statistics/Statistic-on-International-Development-2010/SID-2010-Section-3-How-much-is-UK-expenditure-on-International-Development/ and www.dfid.gov.uk/About-us/How-we-measure-progress/Aid-Statistics/Statistic-on-International-Development-2010/SID-2010-Key-statistics/

Figure 3: Management breakdown of EC external assistance funds in 2010



Source: http://ec.europa.eu/europeaid/files/publications/europeaid_annual_report_2011_en.pdf. EC' refers to the EU institutions, 'EU' refers to the EU institutions plus member states, 'EC budget' refers to external assistance funded via the European Commission budget, and 'EDF' refers to the European Development Fund.

Figure 2: Top 20 donors by ODA commitments in 2010



Source: CRS commitments current from http://stats.oecd.org

Annex 2: Results tables of recent donor quality rankings

Easterly, Williamson (2011): Ranking of donor agencies 2008

Average percentile ranking (higher rank means better practice)

Donor	Rank of overall percent rank	Specialisation	Selectivity	Ineffective channels	Overhead	Transparency	Average of percent rank
Bilateral agency							
United Kingdom	5	34%	80%	70%	83%	82%	70%
Japan	6	54%	71%	41%	95%	57%	63%
New Zealand	8	41%	95%	63%	44%	66%	62%
Germany	9	51%	44%	59%	98%	57%	62%
Ireland	11	44%	61%	85%	78%	36%	61%
Australia	12	56%	78%	4%	90%	66%	59%
Netherlands	13	17%	32%	67%	93%	82%	58%
Luxembourg	14	34%	85%	81%	17%	61%	56%
Norway	15	24%	37%	89%	100%	20%	54%
European Commission	16	39%	83%	22%	59%	66%	54%
Denmark	19	44%	66%	74%	39%	36%	52%
Italy	22	80%	2%	56%	73%	36%	50%
Canada	23	27%	51%	7%	76%	82%	49%
Austria	24	95%	0%	48%	32%	66%	48%
France	26	17%	41%	44%	46%	82%	46%
United States	28	44%	12%	37%	54%	80%	45%
Portugal	29	68%	34%	0%	85%	36%	45%
Sweden	30	12%	20%	93%	61%	27%	42%
Switzerland	32	17%	22%	78%	49%	36%	40%
Spain	34	10%	49%	19%	80%	36%	39%
Finland	36	12%	68%	15%	71%	20%	37%
Belgium	37	29%	27%	33%	41%	55%	37%
Greece	39	61%	7%	11%	27%	36%	28%
Average	**21**	**40%**	**46%**	**48%**	**65%**	**53%**	**50%**
Multilateral agency							
Nordic Development Fund	1	85%	76%		88%	66%	79%
Global Fund	2	88%	88%		68%	64%	77%
Asian Development Bank	3	73%	90%		51%	82%	74%
African Development Bank	4	78%	59%		66%	82%	71%
IDA (World Bank)	10	7%	100%		56%	82%	61%
IDB	18	63%	15%	52%	63%	66%	52%
IMF (SAF, ESAF, PRGF)	20	66%	93%		12%	30%	50%
EBRD	21	83%	10%		24%	82%	50%
CariBank	31	71%	39%		22%	36%	42%
GEF	40	29%	46%		5%	9%	22%
Average	**15**	**64%**	**61%**	**52%**	**46%**	**60%**	**58%**
UN agency							
UNRWA	7	93%	98%		37%	25%	63%
UNFPA	17	100%	73%		7%	30%	52%
UNDP	25	76%	63%	96%	2%	2%	48%
UNICEF	27	59%	24%	100%	34%	14%	46%
IFAD (UN)	33	2%	54%		20%	82%	39%
WFP	35	98%	56%	26%	0%	14%	39%
UNAIDS	38	90%	5%		15%	14%	31%
UNHCR	41	5%	29%		10%	30%	18%
UNTA	42	0%	17%	26%		2%	11%
Average	**29**	**58%**	**47%**	**62%**	**16%**	**23%**	**39%**

Knack et al. (2010): Donor scores and rankings (by sub-index and overall, for 2007)

Donor name	Sub-indexes								Overall index		Avg rank difference of sub-indexes	No. of indicators with data
	Selectivity		Alignment		Harmonisation		Specialisation					
	Score	Rank	Score	Rank	Score	Rank	Score	Rank	Score	Rank		
African Development Bank	0.505	11	-0.359	28	-0.193	26	1.095	5	0.310	13	14.0	15
Asian Development Bank	1.837	2	0.800	2	0.305	14	2.382	1	1.784	1	6.5	15
Australia	0.122	13	-0.042	22	0.708	5	0.253	8	0.348	12	9.3	18
Austria	-0.307	21	-0.654	33	-0.484	31	-0.317	28	-0.641	34	6.5	18
Belgium	0.104	14	0.161	17	-0.191	25	-0.416	29	-0.130	21	8.8	18
Canada	-0.357	23	-0.241	26	-0.416	29	-0.187	21	-0.421	29	4.5	18
Czech Republic	-0.727	28	0.754	36	-1.150	36	-1.016	38	-1.283	37	5.0	14
Denmark	1.158	5	0.711	4	1.015	2	0.064	14	0.985	3	6.2	18
EBRD	0.960	6	-0.077	23	0.129	18	-0.555	34	0.083	16	14.8	15
EC	-0.612	26	-0.466	31	0.384	13	0.212	10	-0.160	23	12.7	17
Finland	0.035	17	0.447	9	0.574	9	-0.419	30	0.233	15	11.8	18
France	-0.908	33	0.180	15	-0.200	27	0.116	12	-0.226	24	12.5	18
GAVI Alliance	-	-	0.752	3	0.862	3	0.145	11	0.861	5	5.3	9
Germany	-0.452	24	0.407	11	0.205	17	-0.198	22	0.026	17	7.3	18
Global Fund	-0.652	27	0.213	14	-0.094	22	0.019	16	-0.137	22	7.5	15
Greece	-0.828	31	-0.535	32	-0.859	34	-0.602	35	-0.975	35	2.3	18
Hungary	-0.753	30	-1.623	38	-0.604	32	1.492	2	-0.558	31	18.3	17
IDB	0.007	18	0.044	20	0.263	16	0.873	6	0.412	9	7.3	15
IFAD	0.886	7	0.676	5	0.669	6	0.232	9	0.831	7	2.2	15
IMF	1.161	4	0.099	18	0.089	19	1.114	4	0.787	8	9.8	12
Ireland	0.099	15	1.453	1	1.145	1	-0.238	23	0.919	4	13.3	17
Italy	-0.486	25	-0.403	29	-0.159	23	-0.115	20	-0.406	28	4.8	18
Japan	-0.349	22	0.475	8	-0.894	35	-0.278	24	-0.333	27	13.8	18
Korea, Republic of	-1.128	36	-0.035	21	-1.409	37	-0.768	36	-1.114	36	8.0	15
Luxembourg	0.675	9	-0.427	30	0.552	10	-0.517	33	0.024	18	15.3	18
Netherlands	1.332	3	0.655	6	0.605	8	-0.022	18	0.838	6	7.8	18
New Zealand	0.573	10	0.070	19	-0.459	30	-0.311	27	-0.090	20	11.3	18
Norway	0.380	12	0.387	12	0.280	15	-0.298	25	0.247	14	7.0	18
Poland	-1.098	35	0.161	16	-0.843	33	1.441	3	-0.046	19	18.8	14
Portugal	-1.018	34	-1.562	37	-2.208	38	-0.044	19	-1.718	38	10.0	17
Spain	0.086	16	-0.289	27	-0.066	21	-0.308	26	-0.233	25	6.3	18
Sweden	0.691	8	0.356	13	0.061	20	0.088	13	0.383	11	6.0	18
Switzerland	-0.063	19	-0.122	25	-0.187	24	-0.495	32	-0.317	26	6.7	18
Turkey	-1.447	37	-0.101	24	0.388	12	-0.887	37	-0.640	33	14.7	13
United Kingdom	-0.143	20	0.421	10	0.765	4	0.023	15	0.396	10	8.8	18
United Nations	-0.833	32	-0.700	34	0.621	7	-0.484	31	-0.479	30	13.7	15
United States	-0.751	29	-0.731	35	-0.331	28	-0.004	17	-0.634	32	9.2	18
World Bank	2.301	1	0.600	7	0.522	11	0.623	7	1.291	2	5.0	15

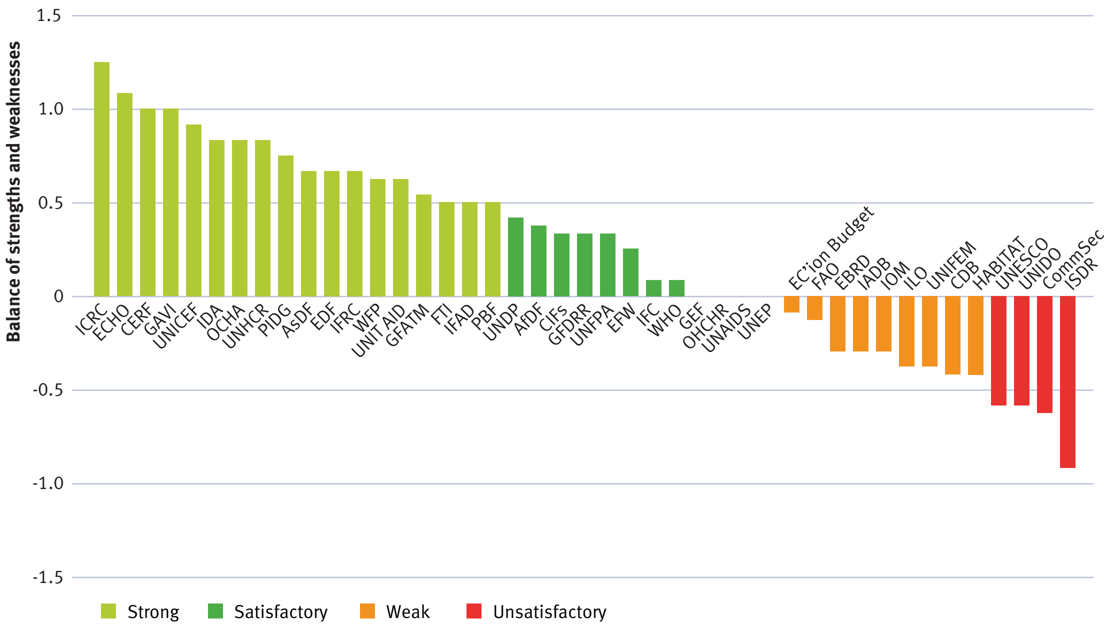
Note: Sub-index scores are normalised in standard deviation units with a mean of 0.

Birdsall et al. (2011): Ranking of donors by aid quality dimensions

Donor	Maximising efficiency	Fostering institutions	Reducing burden	Transparency and learning
Australia	21	19	14	1
Austria	25	29	29	14
Belgium	18	26	21	18
Canada	23	21	23	17
Denmark	13	5	10	4
Finland	16	10	5	12
France	20	17	17	21
Germany	26	14	18	11
Greece	29	31	22	26
Ireland	7	1	7	3
Italy	19	22	19	27
Japan	14	13	27	20
Korea, Republic of	30	16	31	30
Luxembourg	8	27	15	22
Netherlands	17	6	6	9
New Zealand	10	23	13	6
Norway	24	9	20	7
Portugal	6	25	30	19
Spain	31	15	25	8
Sweden	22	11	16	15
Switzerland	26	24	26	26
United Kingdom	12	7	4	13
United States	27	30	28	24
AfDF	2	4	12	25
AsDF	3	3	10	29
EC	11	12	9	2
Global Fund	1	18	11	10
IDA	9	2	2	5
IDB Special Fund	5	8	3	31
IFAD	4	20	1	23
UN (select agencies)*	15	28	24	16

* An aggregation of five UN agenciesused primarily for country-level analysis: the Joint United Nations Programme on HIV/AIDS, the United Nations Children's Fund, the United Nations Development Programme, the United Nations Population Fund, and the World Food Programme.
Source: Birdsall et al. (2011)

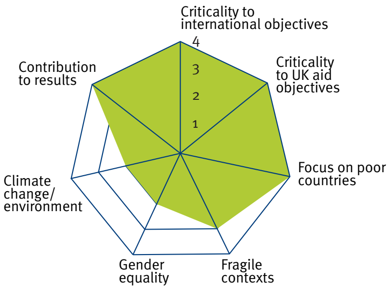
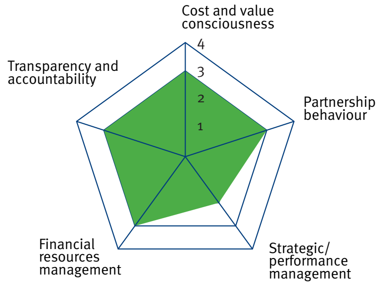
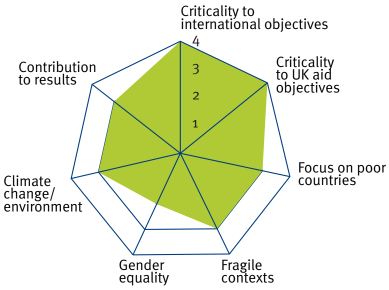
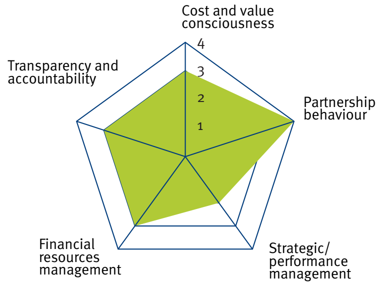
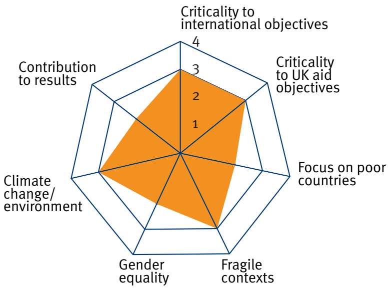
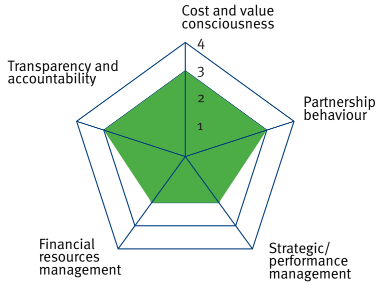
MAR (2011): Contribution to UK development objectives



Note: Where organisations have the same score they are shown in alphabetical order.

MAR (2011): Organisational strengths

Note: Where organisations have the same score they are shown in alphabetical order.

	Contribution to UK development objectives	Organisational strengths	Strength and areas of reform
ECHO Very good value for money			ECHO plays a critically important role. It is strong on delivery and demonstrates good partnership behaviour. It needs to apply value for money considerations more consistently, strengthen the link between humanitarian aid and longer term development, and pay greater attention to gender and environment issues.
EDF Very good value for money			EDF is critical to UK development objectives and achieving the MDGs with its size, poverty focus and unique partnership model. It is committed to transparency and has robust inancial management. However despite strong strategies, procedures remain inlexible and more work is needed to systematically demonstrate delivery against a results framework
European Commission Budget Adequate value for money			The size and reach of EC Budget Instruments is signiicant for UK development objectives. They are transparent, with robust inancial management and strong in fragile states. But a signiicant share of resources goes to Middle Income Countries and focus on MDGs is mixed. Procedures are inlexible and they need to more systematically demonstrate delivery against a results framework.

Overseas Development Institute, 111 Westminster Bridge Road, London SE1 7JD, Tel: +44 (0)20 7922 0300, Email: publications@odi.org.uk. This and other ODI Background Notes are available from www.odi.org.uk.

 ISSN 1756-7610.