

Adjusting indices of donor aid quality

By Matthew Geddes

The contribution of aid to development depends on both the volume of resources and on the effectiveness of delivery. Three recent exercises ranking the ‘quality’ of donors (Easterly and Williamson, 2011; Knack et al., 2011; Birdsall et al., 2010) recognise this and have produced indices of donor aid quality (Box 1). They reflect a resurgence of the comparative analysis of donor practice, focussing on aid quality issues, value for money, and transparency. Their evidence is a valuable starting point for discussions of donor practices, and all three stimulate comparative learning and peer pressure as a tool for improved practice.

Similar index methodologies are used by donors for aid allocation decisions, including the recent multilateral aid review by the UK Department for International Development (DFID), with direct influence on the spending of millions of dollars of aid. As more use is made of these methods, their results and the questions they raise increase in importance.

This Background Note suggests three potential adjustments to improve the quality of the evidence:

- The three exercises combine indicators representing ‘international best practice’ as a proxy for the impact of aid. Recipient perspectives would augment this approach and provide a more direct measure of impact.
- The impact of an aid delivery approach is context specific – a practice that is effective in one country may not be in another – but none of the exercises reflect this. Measuring the performance of donors in each aid recipient country relative to others would isolate donor specific factors that

Box 1: Three indices of donor aid quality

Easterly and Williamson (2011) measure donors’ adherence to best practices (measures of aid quality) as defined by aid agencies, outside aid monitors and academic literature. They acknowledge no direct evidence of links to aid impact. The five best practice dimensions are listed as: transparency; overhead costs; specialisation/fragmentation; use of effective channels; and allocation to less corrupt, more democratic, or more free recipients, with multiple indicators for each. The absolute performance of donors is measured, including trends over time and using equal weightings.

Birdsall et al. (2011) investigate aid agency effectiveness (aid quality) concentrating on the measures that aid agencies control. Thirty indicators are used as part of four pillars of aid quality (Maximising efficiency; Fostering institutions; Reducing burden; Transparency and learning) based on the academic literature and consensus in the development community. Most donors are revealed to have strengths and weaknesses, although some perform above or below average on most indicators.

Knack et al. (2011) construct an aid quality index using 18 indicators drawn from the OECD Paris Declaration Monitoring Survey and OECD Development Assistance Committee (DAC) tables and aid data, grouped into four sub-indices: selectivity, alignment, harmonisation and specialisation. The impact of changed weightings on overall rankings is highlighted, with conclusions on whether a single index of donor quality is valid and reliable. Factors beyond each donor’s control are adjusted for.

affect aid quality and reflect the context of aid delivery when measuring effectiveness.

- The exercises focus on aid quality but this misses the contribution of volume of aid to impact. When donors’ scores are aggregated across recipient

countries, the score for a donor with many low volume country programmes will be skewed. Once scored on aid quality, donor performance in each recipient country can be weighted by the amount of aid a donor supplies to that country, so the aggregate figure represents the average quality of aid delivered.

The Background Note concludes by linking these adjustments to underlying theoretical and methodological problems of indices before discussing how the results are best interpreted.

Adjustment 1: Incorporating aid recipient perspectives

If we are going to have a beauty contest, then at least we should let the audience vote. The best measure of aid quality is its impact. However, lack of comparable and disaggregated data on aid impact leads the three exercises to follow a more indirect approach, using intermediate indicators from the literature to proxy for high quality aid outcomes.

There are problems with this approach. There is little evidence showing that the indirect measures of aid quality used in the exercises lead to greater aid impact for recipients. The Paris Declaration Evaluation limits itself to suggesting a ‘plausible contribution’: certainly not a necessary, sufficient or direct link. In addition, the available proxies are often poor and lacking data, making it difficult to compare donors over time (Box 2).

One alternative that bridges this knowledge gap is to develop measures of aid quality based on aid recipients’ direct experience of the impact – both positive and negative – of the intervention.

- Recipient perspectives can accommodate data issues, such as country specifics, where in-depth knowledge is needed to attribute impact. Recipient perspectives provide better coverage and allow expansion of data. They are not limited to DAC members and can distinguish between modalities/agencies as well as noting less quantifiable aspects of aid practice.
- The feedback loop would be less driven by donor perspectives, making aid more accountable

Box 2: Known issues with proxy indicators of donor quality

- **Many aspects of aid quality are not easily measured.** The example given in Knack et al. (2010) is specialisation. To have a positive impact, the area of specialisation needs to tally with the comparative advantage of each donor, on which there are no data. Instead, the exercises assume – without evidence – that any specialisation follows each donor’s comparative advantage. If data were available and included, the rankings would change. So, is what is included a sufficient proxy for aid quality?
- **External factors are overlooked.** While alignment and harmonisation are desirable, they provide no guidance on such questions as what donors should do if the government does not take the lead on development. Aligning to a government with no interest in poverty reduction is unlikely to reduce poverty (Booth, 2010).
- Aid effectiveness often focuses on lessons about ‘what not to do’. To create positive indicators, the exercises often reverse the logic. However, while the evidence suggests that high levels of technical cooperation are inefficient, this doesn’t mean that zero technical cooperation would be better. A linked issue is that most of the indicators are interpreted in a linear fashion, while many may have a non-linear impact. **The optimum level for most indicators is unclear** and is affected by thresholds, country circumstances and other constraints, particularly those that overlap.
- **It is unclear whether the ‘best practices’ form a cohesive blueprint** – do they work well together or are there trade-offs between them? Lower administrative costs are positive for one aspect of aid quality but negative for another: the cost savings may reduce the number of staff for disbursement in the poorest countries where government capacity is weaker. So, it is not certain that aid quality measures suit aggregation into a single index. All three exercises caution against overall rankings, indicating that a separate ranking for each dimension allows users to identify the specific strengths and weaknesses of each donor. Easterly and Williamson (2011) find little correlation between sub-components in their index, suggesting that these are independent measures of aid quality.
- Although discussed in Knack et al. (2011), **none of the rankings reflect changes in optimal aid quality behaviour among donors.** If all donors crowded into the neediest country or sector the global impact would be undesirable, even if it improved their individual effectiveness.
- Many indicators do not apply to increasingly diverse development finance options outside official development assistance (ODA). Their trade-offs **may not work well for climate finance, blended private sector support to middle income countries (MICs) or South-South cooperation** provided as technical assistance. Do they even apply to humanitarian aid (Darvill, 2011)?
- **Some key aspects for aid recipients are not addressed:** the translation of pledges into commitments, the importance of predictability by financial quarter, mutual respect and donors’ ability to adapt programmes to country contexts (Wathne and Hedger, 2009). Even previously included indicators such as the generosity of repayment terms are now excluded.

to those it aims to assist: a recognition that recipients have different knowledge, beliefs and priorities on aid, yet no vote to communicate dissatisfaction (Easterly and Williamson, 2011; Barder, 2010).

One approach would be to have aid recipients choose relevant indicators and provide raw data for the exercises. Successful examples include the Voices of the Poor survey (World Bank, 1999), which gathered the thoughts of over 40,000 people, and the Humanitarian Response Index (HRI, 2011), which surveyed over 2,000 respondents, recording their perceptions of the donors with whom they work.

Other examples include the Listening Project (2011), which gathered opinions from aid recipients on the five Paris declaration principles, and Wathne (2009) who focused on additional indicators to expand the Paris Declaration based on surveys of recipient priorities. Further examples are included in the additional resources listed at the end of this Background Note.

The analysis in the three exercises is complementary to the work on recipient perspectives. A specific survey of recipient perspectives to provide both data and indicator selection information would be a worthwhile effort if we are serious about monitoring aid quality.

Melamed (2011) discusses why we are so bad at translating the views and priorities of poor people into development decision-making. While there may be methodological and logistical barriers, such as selecting who represents recipients and how to combine their preferences, the potential gains are enormous for an industry that spends between \$150 and \$250 billion a year. This study can't cover that process but it can demonstrate similar effects using a simpler approach.

Adjusting the weightings of the indicators used by the three exercises according to aid recipient priorities would allow recipients to increase the value of indicators they see as important for aid impact and devalue those seen as less important. The weights matter, because where the indicators that proxy for aid quality are not correlated (common in the three exercises), the results are sensitive to the choice of weights. These weights embody, therefore, the trade-offs between different aspects of aid quality.

Birdsall et al. (2010) say that the lack of alternative approaches leads the researchers to weight the indicators equally when they are aggregated. Yet it is unlikely that each indicator (or cluster) is of equal importance. In the (simpler) methodology to include recipient perspectives demonstrated below, recipient priorities define the weights that they feel maximise the impact of aid. These weights will differ among aid recipient countries and donors.

The illustration below considers Viet Nam and indicators for the numbers of Parallel Implementation Units (PIUs) set up by donors, Donor Missions and Coordinated Analytical Works. Data are from the 2011 Paris Declaration Monitoring Survey (OECD, 2011). Table 1 shows raw data in the first three columns, rankings in the middle three columns and the results of equal weighting in the last two columns. Finland is the highest ranked donor, with its highest aggregated score of 4.7 across the three indicators.

Donors were then re-ranked using illustrative aid recipient preferences, making Number of Missions the least important indicator, Parallel PIUs twice as important and Coordinated Works three times as important. As a result, donors like the EU Institutions and the World Bank, which score poorly on Number of Missions, are seven and nine places higher respectively when the weightings reduce the contribution of that indicator.

Table 1: Raw data and equal weighting scores (top 10 in descending order by Equal Weighting Rank)

Donor	Raw Data			Equal Weighting Indicator Rank			Equal Weighting Final Score	
	Parallel PIUs	Number of missions	Coordinated Works	Parallel PIUs	Number of Missions	Coordinated Works	Aggregation	Equal Weighting Rank
Finland	0	5	6	1	7	6	4.7	1
Sweden	0	7	9	1	11	3	5.0	2
United Kingdom	0	9	12	1	12	2	5.0	2
Hungary	0	0	0	1	1	16	6.0	4
Spain	0	0	0	1	1	16	6.0	4
Switzerland	0	0	0	1	1	16	6.0	4
United States	0	5	4	1	7	10	6.0	4
Australia	0	11	7	1	13	5	6.3	8
New Zealand	0	5	3	1	7	11	6.3	8
GAVI Alliance	0	2	0	1	6	16	7.7	10

Table 2: Adjusted weightings and comparison (top 10 in descending order by Recipient Weighting Rank)

Donor	Recipient Weighting Indicator Ranks			Recipient Weighting Final Score		Comparison		
	Parallel PIUs	Number of missions	Coordinated Works	Aggregation	Recipient Weighting Rank	Equal weighting	Recipient Weighting	Reward / Penalty
United Kingdom	1	6	3	10	1	2	1	1
Sweden	1	5.5	4.5	11	2	2	2	0
Finland	1	3.5	9	13.5	3	1	3	-2
Australia	1	6.5	7.5	15	4	8	4	4
EU Institutions	1	8	10.5	19.5	5	12	5	7
United States	1	3.5	15	19.5	5	4	5	-1
New Zealand	1	3.5	16.5	21	7	8	7	1
World Bank	1	13.5	10.5	25	8	17	8	9
Hungary	1	0.5	24	25.5	9	4	9	-5
Spain	1	0.5	24	25.5	9	4	9	-5

Table 2 compares the recipient weighting with equal weighting. As expected, re-ranking the donors changes our understanding of which has provided the aid of highest quality. The shifts could be even more dramatic with a full process of aid recipients guiding choice of indicators and supplying a share of the underlying data.

Neither recipient nor researcher views can capture all aspects of aid quality. However, a weighting that includes aid recipients' experience and preferences will add value. The equal weighting used in the three exercises is not neutral and good alternatives exist (see Lugo and Decanq, 2011).

By extension, each type of stakeholder surveyed may have a different viewpoint on the most important factors driving the quality of aid in their country, as shown by such country level aid effectiveness frameworks as the Rwandan Donor Performance Assessment Framework. Although requiring far greater effort, this would allow the rankings to better approximate the true complexity of aid quality.

Adjustment 2: Allowing for country context

The exercises have been criticised for penalising donors that, by choice or mandate, work in countries where it is difficult or inappropriate to provide aid on the basis of the 'internationally recognised best practices'. For example, donors working in fragile states, with additional security concerns and weak government systems, may be penalised by the exercises.

For Easterly and Williamson (2011), conditional measures are better. But, given the impossibility of controlling for all relevant factors (some not even measurable), they feel it is better not to control for any, as controlling selectively introduces bias. Knack et al. (2011) control for recipient country

effects using regression analysis. They find that, in some cases, recipient characteristics matter more for the score than donor characteristics.

An alternative approach is to use the fact that all donors confront the same factors in each recipient country. This means controlling for external factors by evaluating the performance of donors relative to the performance of all other donors in that country. This also captures factors for which there are no data because they are immeasurable or not measured.

As the practices required to deliver high quality aid vary in each recipient country so will the definition of aid quality. This seems to fit better with the original rationale for measuring donor performance i.e. abstracting from the countries where they work and, instead, focussing on whether each donor delivers high quality aid that is appropriate to the unique situation of each aid recipient country.

The same donor behaviour means different things in different contexts. For example, the use of country systems at a level of 25% in Mauritania may be a success, but would be considered a failure next door in Morocco. Typically, however, the exercises treat both scores as equal. Similarly, when working in fragile states, different aid delivery behaviours from those used in non-fragile countries are required to maintain aid quality. The delivery approach that maximises positive impact may vary among countries.

Table 3 uses data on the share of aid using country procurement systems in Sierra Leone and Rwanda, taken from the 2011 Paris Declaration Monitoring Survey and showing their standard and relative scores. In reality, scores would be aggregated across multiple indicators and countries. For clarity, just one indicator is presented here, with only donors that give aid to both countries included. With such a limited example, little change is expected in the ranking.

However, by standardising a donor's score around the country mean it is possible to see it in relation to the country average. The mean score for Sierra Leone is 19%, and for Rwanda it is 60%.

This system rewards donors that make the greatest effort to deliver high quality aid in the countries in which they work, rather than rewarding countries that deliver high quality aid by working in 'easy' countries or penalising donors that are constrained by their mandate.

This approach fits two key ideas. First, that institutional commitment to aid quality is fairly uniform but that every donor has a range of aid delivery practices in different countries. Second, aid quality is country specific and best assessed at the country level.

In the standardised scores (which adjust for the country context) a donor that supports Sierra Leone is rewarded for delivering support at a higher level of quality than the average donor and is not penalised for supporting a country where difficulties in delivering high quality aid keep the average aid quality low. Looking at the figures for the EU Institutions, in the normal ranking, the 37% score in Sierra Leone is

ranked lower than the score of 85% in Rwanda. In the standardised ranking, the opposite is true as the ranking adjusts for the lower average use of country procurement systems in Sierra Leone. The Sierra Leone score is higher than the Rwanda score as it is even further above the country average. Similarly, the Global Fund is rewarded far more for its score of 100% in Sierra Leone than for its score of 96% in Rwanda.

Adjustment 3: Accounting for aid volume across countries

The three exercises aggregate scores across all relevant aid recipient countries to get an overall donor score. While the aggregate score does not reflect or relate to donor practice in any one country, it also means that a donor with many small, high quality aid programmes and a few much larger low performing aid programmes will receive a high score, even though most of their aid (by volume) is of lower quality.

Table 4 shows the effect of weighting donor scores by the volume of their aid to several recipient countries (see the similar approach in the Commitment

Table 3: Comparison between standard country scores and scores relative to the country mean

Donor	Normal				Standardised				Change in rank
	Sierra Leone	Rwanda	Score	Rank	Sierra Leone	Rwanda	Score	Rank	
United Kingdom	5%	99%	52%	3	-0.66	1.09	0.21	4	-1
GAVI Alliance	0%	0%	0%	8	-0.89	-1.45	-1.17	8	0
Global Fund	100%	96%	98%	1	3.62	1.01	2.32	1	0
United States	0%	0%	0%	8	-0.89	-1.45	-1.17	8	0
Germany	7%	74%	40%	5	-0.56	0.44	-0.06	5	0
Japan	0%	44%	22%	7	-0.89	-0.32	-0.61	7	0
World Bank	25%	79%	52%	4	0.24	0.57	0.40	3	1
EU Institutions	37%	85%	61%	2	0.76	0.74	0.75	2	0
United Nations	1%	60%	30%	6	-0.87	0.10	-0.38	6	0

Table 4: Change in donor rank when scores are adjusted for the volume of aid (top 10 by unweighted average)

Donor	Share of aid to government using country procurement systems		Unweighted Average	Rank	ODA Volume USD Millions		Weighted Average	Rank	Change in rank
	Viet Nam	Zambia			Viet Nam	Zambia			
	IFAD	100%			100%	100%			
Netherlands	99%	99%	99%	2	32	13	99%	2	0
Ireland	94%	100%	97%	3	20	9	96%	3	0
Denmark	100%	94%	97%	4	21	47	96%	5	-1
United Kingdom	92%	100%	96%	5	63	58	96%	4	1
World Bank	72%	69%	70%	6	39	1680	69%	7	-1
Germany	74%	55%	64%	7	27	47	62%	9	-2
Finland	68%	56%	62%	8	16	18	62%	8	0
Japan	4%	90%	47%	9	37	1068	87%	6	3
Sweden	88%	0%	44%	10	13	12	46%	11	-1

to Development Index). The example is based on the share of each donor's aid that uses country procurement systems in Viet Nam and Zambia. Including only donors that give aid in both countries, it weights the scores by the volume of aid provided by the donor as a share of that donor's total aid.

Weighting by volume affects individual scores. Japan makes greater use of country procurement systems in Zambia (90%), where it supplies a proportionally larger volume of aid, than in Viet Nam (4%). When the aggregation takes aid volume into account, Japan's ranking increases by three places. While changes on the basis of just two countries are relatively small, they may be significant over a large number of countries.

If donors provide less aid volume where conditions are most difficult because the expected impact is lower, this methodology guards against punishing donors for working in more challenging countries.

Interpretation and transparency

The first two adjustments seek to mitigate issues caused by the use of proxy indicators to bridge the knowledge gap between aid practices and aid impact. However, many serious problems remain (Box 2), making it difficult to say whether a donor's ranking in the three exercises is a good indicator of the impact that donor has on the ground. At best, it is a starting point for issues to investigate further.

Knack et al. (2011) state that 'most indicators of donor performance are based on plausible but largely untested beliefs about best practices in aid management.' Wathne and Hedger (2009) find some donor procedures where there is no consensus on what constitutes best practice and many principles that need to be adapted to become appropriate for different countries. This mirrors the global situation where the Paris Declaration is less of a consensus in practice than in presentation, with few donors straining themselves to meet the international targets.

The resulting scarcity of evidence means that, despite similar aims, there is limited overlap between the three exercises in terms of indicators and they produce different rankings for the same donors. Which index to believe? The answer is that none of them actually claim to measure aid quality. While they measure adherence (to a set of indicators) there is insufficient evidence to claim that they measure performance.

The role of the researchers is to demonstrate that adherence to the chosen set of indicators will lead to the best performance, given the prevailing context. The debate over quality, therefore, centres on the ability of the researchers to support their chosen sets of indicators with country level evidence incorporat-

ing the context in which aid is given. The difficulty, from a researcher's perspective, is why the suggestions above of incorporating recipient perspectives and assessing donors in relation to their peers are valuable tools.

Sensibly, all three exercises caution against over-interpretation of their results. Easterly and Williamson (2011) state that their ranking reflects scores on five good practices, not whether a particular agency's aid is effective at achieving good results. Knack et al. (2010) suggest that, because of alternative plausible assumptions on the weights or the indicators, the reader should interpret their rankings as a general guide and not draw strong conclusions from small differences between donors. While it is certainly useful to know that donor X is a reasonable choice if it appears in the top half of the rankings, it is unclear whether this is sufficient for a bilateral donor to allocate scarce resources amongst competing multilateral partners.

Given the lack of evidence or agreement, transparency is essential as all the choices and adjustments made by the researchers influence the final scores. All three exercises should be commended for their transparency. By avoiding a 'black box' scenario and making their judgements explicit and the raw data available online, or by using publicly available data, they allow third parties to replicate the methods so that alternative weightings, indicators and trade-offs can be tried and compared. Impressively, the website for Birdsall et al. (2011) (www.cgdev.org/QuODA) even allows readers to enter their own weightings. A common format would make this process even easier and allow tests, such as using data from one exercise with the aggregation methodology of another.

Conclusion

Whereas previous decades of development have boasted about the expanding volume of aid, the financial crisis afflicting many major donors in 2012 makes a repetition of this unlikely. Efforts to improve the development impact of many donors' efforts must, therefore, focus on improved aid quality. This Background Note suggests three adjustments to existing donor quality exercises, emerging from three straightforward ideas in development practice:

- Aid recipients have valuable insights into the development processes that affect them and their views are an important tool to measure the quality of aid.
- Aid quality is an issue that, for most purposes, is best analysed at the recipient country level, given the importance of country specific factors.
- The volume of aid delivered at each level of quality affects the eventual impact.

The examples shown change the rankings of the donors and demonstrate that the methodologies used by these exercises have a direct impact on the results that are produced. They also show that the indices would benefit from being backed by indicator and weighting choices based on country specific work.

These issues also highlight the importance of a shift in those driving the aid effectiveness debate from donors and donor country researchers to aid recipient governments and populations. This shift can be taken much further (Box 3).

Additional work is required to take this agenda further. Methodologies should be developed to collate aid recipient perspectives and these should be expanded to go beyond government perspectives. Moving the analysis down to the national level is an important step but this also highlights the huge potential insights from further breakdowns by, for example, sector and by the different aid modalities. After this, one key question to ask is ‘why do donor performances differ?’ and then link the insights from the indices to the motivations for aid and the political challenges to achieving best practice in donor countries.

The demand already exists for recipient perspectives, the importance of country context is well known in field offices, similar methodologies are already implemented in other contexts, and the tools are well within current capabilities. The next step is to bring these strands together at both the recipient country and global levels.

Box 3: Shifting the aid effectiveness debate south

Aid recipients already have notions of aid quality that take into account the different competencies and institutional constraints associated with each donor. Donors with domestic politics that do not favour the use of country systems may be guided towards investing in large scale infrastructure, while those better at integrating with national systems are guided towards a service delivery role.

Aid recipients are also concerned with balancing quality against quantity: accepting more aid of lower quality is often a smart choice for them. So, while academic, international and donor perspectives aim to identify and monitor a more singular notion of aid quality, this may not even be an objective for aid recipient governments. They aim to assign tasks and roles in national development to different donors that match their skills and institutional constraints, and all within a framework that provides the government with the information they need to perform this role.

Recipient perspectives will improve the link between aid quality and aid impact, relative performance measures will focus donors on the opportunities to deliver higher quality aid while taking account of the country context, and adjusting for volume will allow the rankings to reflect more accurately the situation on the ground. Finally, increased clarity will make the resulting indices easier to interpret and use properly – essential given their impact on comparative donor practice and spending allocation.

Written by Matthew Geddes, ODI Research Officer (m.geddes@odi.org.uk). Thanks are due to the UK Department for International Development (DFID) and to AusAID for funding this work, and to Steve Knack, Halsey Rogers and Sam Moon for their comments.

References and endnotes

References

- 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
- Booth, D. (2010) 'Country Ownership' when there is no Social Contract: Towards a Realistic Perspective'. Lecture, 13 December 2010. SID-NL Lecture Series 2010-2011. http://sidnl.files.wordpress.com/2010/11/summary-sid-lecture-david-booth_final2.pdf
- CGD (2011) The Commitment to Development Index. Washington: Centre for Global Development (www.cgdev.org/section/initiatives/_active/cdi/)
- Darvill, S. (2011) Humanitarians and the road to Busan. ALNAP blog: <http://www.alnap.org/blog/47.aspx>
- DFID (2011) 'Multilateral Aid Review'. London: Department for International Development. http://www.dfid.gov.uk/Documents/publications1/mar/multilateral_aid_review.pdf
- 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>
- 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>
- Lugo, M. and Decancq, K. (2011) 'Weights in multidimensional indices of well-being: An overview'. Washington DC: World Bank. http://worldbank.academia.edu/mariaanalugo/Papers/271422/Weights_in_multidimensional_indices_of_well-being_An_overview
- Melamed, C. (2011) 'Does development give poor people what they want?'. London: Overseas Development Institute. <http://www.odi.org.uk/resources/details.asp?id=5360&title=poor-people-poverty-development-qaly-value-money>
- OECD (2011) 'Aid Effectiveness 2005-10: Progress in Implementing the Paris Declaration'. Paris: OECD. <http://www.oecd.org/dataoecd/25/30/48742718.pdf>
- Wathne, C. and Hedger, E. (2009) 'Aid effectiveness through the recipient lens'. London: Overseas Development Institute. <http://www.odi.org.uk/resources/download/2746.pdf>
- 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>
- World Bank (1999) 'Voices of the Poor'. Washington DC: World Bank. <http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTPOVERTY/0,,contentMDK:20613045~isCURL:Y~menuPK:336998~pagePK:148956~piPK:216618~theSitePK:336992,00.html>
- DRI (2004) 'The Effectiveness of Aid to Africa Since the HIPC Initiative: Issues, Evidence and Possible Areas for Action'. London: Debt Relief International. http://www.hipc-cbp.org/files/en/open/Advocacy/DRIdocs/DFI_Aid_Effectiveness.pdf
- Gallup (2009) 'Sub-Saharan Africans Rank the Millennium Goals'. Washington DC: Gallup. <http://www.gallup.com/poll/116116/sub-saharan-africans-rank-mdgs.aspx>
- Government of Rwanda (2010) 'Rwanda Donor Performance Assessment Framework' Kigali: Government of Rwanda. http://devpartners.gov.rw/index.php?option=com_content&task=view&id=61 and for the detailed scores <http://www.africa-platform.org/download/file/dfd/424>
- HRI (2011) 'Humanitarian Response Index'. Madrid: Dara. <http://daraint.org/humanitarian-response-index/humanitarian-response-index-2010/information-sources/>
- Keystone Accountability (2011) 'NGO Partner Survey 2010'. London: Keystone Accountability. http://www.keystoneaccountability.org/sites/default/files/Keystone%20partner%20survey%20Jan2011_0.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
- Multilateral Operational Performance Assessment Network (2011) 'MOPAN Survey 2011'. Montreal: Universalia. <http://www.mopanonline.org/>
- OECD (2011) 'Partner Country Position Paper for Busan'. Paris: OECD. http://www.aideffectiveness.org/busanhlf4/images/stories/hlf4/Partner_Country_Position_Paper_-_Final.pdf
- Oxfam (2004) 'Paying the Price: Why rich countries must invest now in a war on poverty'. Oxford: Oxfam International. <http://www.oxfam.org/sites/www.oxfam.org/files/price2.pdf>
- Strategic Partnership for Africa (2009) 'Budget Support Alignment Survey'. London: Overseas Development Institute. <http://www.spasurvey.info>

Endnotes

- Such methodological issues require careful thought to collect high quality evidence. The Voices of the Poor approach would be more useful with stand-alone donor projects; less so when donors work through country systems to support country strategies. Here, it is difficult for citizens to judge what individual donors are doing. While evidence from aid recipients has its own biases, these need to be understood and minimised. It is important to manage heterogeneity (when aggregating evidence collected in different ways); adjudicate between perceptions and other forms of evidence, and identify 'gaming' of responses.
- This uses a z-scores methodology similar to the use of z-scores in QuODA to benchmark donors' relative performance across time periods. Z-scores are a standardisation following the formula $z = (x - \mu) / \sigma$, where x is a raw score to be standardised; μ is the mean of the population and σ is the standard deviation of the population. The QuODA team identify similar issues, e.g. an individual donor may have improved performance from one year to the next, but their score may fall if other donors improve by a greater amount. QuODA also demonstrate that it is possible to compute comparable scores across time periods to get around this issue.

Additional resources on aid recipient perspectives

- African Development Bank (2010) 'The Tunis Consensus: Targeting Effective Development'. Tunis: African Development Bank. http://www.aideffectiveness.org/images/tunis/Tunis_Consensus_3mars.pdf
- Aid Thoughts (2011) Comment #3 by Maria Ana <http://aidthoughts.org/?p=2167#comment-1895>
- Barder, O. (2010) What can development policy learn from evolution? Owen abroad: thoughts on development and beyond. Blog: <http://www.owen.org/blog/4018>

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.



Readers are encouraged to reproduce material from ODI Background Notes for their own publications, as long as they are not being sold commercially. As copyright holder, ODI requests due acknowledgement and a copy of the publication. For online use, we ask readers to link to the original resource on the ODI website. The views presented in this paper are those of the authors and do not necessarily represent the views of ODI. © Overseas Development Institute 2012. ISSN 1756-7610.