

A DROP IN THE OCEAN? THE INTERNATIONAL DEVELOPMENT TARGETS AS A BASIS FOR PERFORMANCE MEASUREMENT

Howard White*
Institute of Development Studies
University of Sussex
(h.white@ids.ac.uk)

1. INTRODUCTION

The 1996 OECD document *Shaping the 21st Century* laid out seven quantifiable development targets, which have become known as the International Development Targets (IDTs), or sometimes International Development Goals. Although there have been very many previous development targets over the decades, this set of targets have won unprecedented support and predominance. The UK Department for International Development (DFID), and its Secretary of State, have been particularly vocal in promoting the IDTs. The targets have occupied a central position in the two White Papers, public pronouncements of the Secretary of State, and within DFID in developing its strategy.

This paper examines the use of the IDTs for measuring the performance of development agencies such as DFID. It begins with a brief introduction to the targets and why targets matter. Part 4 discusses desirable properties of performance measures, their use and appraises the IDTs against these requirements. Part 5 looks at the use of the IDTs by DFID and how they have handled some of the problems mentioned in Part 4. Part 6 concludes.

2. BACKGROUND ON THE IDTS

The International Development Targets are listed in Table 1. The targets constitute poverty reduction goals embodying a multi-dimensional conception of poverty (see Baulch, 1996, and White, 1997 for discussion). Poverty is about more than just a lack of income, it includes deprivation with respect to other aspects of well-being such as health and education.

The main targets are quantifiable. The exception is the target for access to reproductive health services, for which there is no corresponding indicator. Contraceptive prevalence is commonly used, though this is problematic since there is no agreed target level partly since contraception is not acceptable in some cultures,¹ and even where it is acceptable the desired level depends on desired fertility. The targets are outcome indicators,² other than those for the environment, reproductive health and education. Accompanying the targets is a statement recognising the importance of qualitative factors related to governance. Although there are numerical indicators for these aspects (such as those produced by Freedom House on political freedom and civil liberties)³ they have not been used to monitor progress on the IDTs since donors have been unable to reach agreement on what should go into such indicators (see below for more discussion on this point).

* Report prepared for the UK National Audit Office (NAO). The opinions are those of the author alone and are not attributable to NAO or any of its employees.

¹ Or at least certain types of reproductive health service are unacceptable. The US aid agency, USAID, is forbidden by law to support programmes which provide abortion services.

² The different classifications of indicators are discussed in more detail below. See also Carvalho and White (1994).

³ Downloadable from www.freedomhouse.org.

Table 1 The International Development Targets as stated in *Shaping the 21st Century* and their provenance

<p><i>Economic well-being</i></p> <p>The proportion of people living in extreme poverty in developing countries should be reduced by at least one half by 2015.</p>	<p>Copenhagen Declaration and Programme of Action 1995</p>
<p><i>Social development:</i></p> <p>There should be substantial progress in primary education, gender equality, basic health care and family planning, as follows:</p> <p>There should be universal primary education in all countries by 2015</p> <p>Progress toward gender equality and the empowerment of women shall be demonstrated by eliminating gender disparity in primary and secondary education by 2015. The death rates for infants and children aged under five years should be reduced in each developing country by two-thirds the 1990 level by 2015. The rate of maternal mortality should be reduced by three quarters during this same period.</p> <p>Access should be available through the primary health care system to reproductive health services for all individuals of appropriate ages, including safe and reliable family planning methods, as soon as possible and no later than the year 2015.</p>	<p>Jomtien Conference on Education for All 1990, and endorsed at Copenhagen Summit on Social Development 1995 and Beijing Conference on Women 1995. Cairo Conference on Population and Development 1994, Beijing and Copenhagen. Cairo. Maternal mortality confirmed at Beijing.</p> <p>Cairo.</p>
<p><i>Environmental sustainability and regeneration</i></p> <p>There should be a current national strategy for sustainable development, in the process of implementation, in every country, by 2005, so as to ensure that current trends in the loss of environmental resources – forests, fisheries, fresh water, climate, soils, biodiversity, stratospheric ozone and the accumulation of hazardous substances and other major indicators – are effectively reversed at both global and national levels</p>	<p>Rio Conference on Environment and Development, 1992.</p>
<p>Source: Development Assistance Committee (DAC), <i>Shaping the Twenty-first Century</i> [Paris: OECD], 1996.</p>	

The targets have not remained fixed. An additional one has been added on HIV/AIDS (a 25 per cent reduction in HIV infection rates among 15-24 year olds in the worst affected countries by 2005 and globally by 2010). And that on the environment has been changed from the requirement for a national strategy for sustainable development to the existence of effective processes for sustainable development. These changes are not universally recognised. Although DFID staff are aware of the new targets (indeed, those in environment are very keen that the new wording be used) the second White Paper used the original list (i.e. without HIV/AIDS and the old wording for the environmental target), though missing the text on the importance of governance.

There are more IDT-indicators than those explicitly mentioned in the targets. A list of indicators has been developed around each target. These are shown in Table 2. Many of these

additional indicators are also outcome indicators. In some cases they are output indicators (e.g. attended births) related to the final outcome given in the IDT. But in the three cases when the target is an output indicator (education, reproductive health and environment) some of the related indicators capture outcomes (literacy, fertility, and, for example, safe water).

Table 2 IDT-related indicators

Target	Indicators
Poverty reduction	Population below a dollar a day Incidence times depth of poverty Poorest fifth's share of national consumption
	Prevalence of underweight under 5
Universal primary education	Net primary enrolment ratio Survival to fifth grade of primary education Literacy rate of adults
Gender equality	Ratio of girls to boys in primary and secondary education Ratio of literate females to males
Infant mortality reduction	Infant mortality rate
Child mortality reduction	Under five mortality rate
Maternal mortality reduction	Maternal mortality ratio Births attended by skilled health personnel
Reproductive health	Contraceptive prevalence rate Total fertility rate
Environment	Existence of national strategies for sustainable development
Sustainable environment	Population with access to safe water Intensity of freshwater use Biodiversity: land area protected Energy efficiency: GDP per unit of energy use Carbon dioxide emissions
Source: DAC 'Methodological Note' DCD/DAC (98)6/ADD.	

The IDTs were adopted at the 34th High Level Meeting of the Development Assistance Committee (DAC) on 6-7 May, 1996 in Paris. There have been comments that it was inappropriate for a developed country group to set targets for developing countries. There are two defences against this criticism. First, it is not necessary for developing country government to sign up to precisely these poverty reduction goals. But donor support is predicated upon recipient commitment to poverty reduction, measured against targets they may choose themselves.⁴ Second, the IDTs are based on resolutions passed at various international conferences (Table 1) and so have already been endorsed by developing countries. In some cases, such as the target for income-poverty, it is the principle which has been adopted at UN conferences rather than the specific target.

⁴ For example, the target of the Ugandan government is to reduce the poverty headcount to 10 per cent or less by 2016, in addition to specific target figures for infant and child mortality.

This discontent manifested itself at the Millennium Summit in New York held from 6-8th September 2000 where an alternative set of development targets, the Millennium Development Goals (MDGs) were adopted (UN, 2000). Whilst similar to the IDTs, they are not the same. They are:⁵

- To halve, by the year 2015, the proportion of the world's people whose income is less than one dollar a day and the proportion of people who suffer from hunger and, by the same date, to halve the proportion of people who are unable to reach or to afford safe drinking water.
- To ensure that, by the same date, children everywhere, boys and girls alike, will be able to complete a full course of primary schooling and that girls and boys will have equal access to all levels of education.
- By the same date, to have reduced maternal mortality by three quarters, and under-five child mortality by two thirds, of their current rates.
- To have, by then, halted, and begun to reverse, the spread of HIV/AIDS, the scourge of malaria and other major diseases that afflict humanity.
- To provide special assistance to children orphaned by HIV/AIDS.
- By 2020, to have achieved a significant improvement in the lives of at least 100 million slum dwellers as proposed in the "Cities Without Slums" initiative.

Despite this criticism, and the emergence of the rival MDGs, there is no doubt that the IDTs have gained a dominant place on the development agenda, being adopted by many development agencies, including the World Bank and the IMF.

Why have the IDTs gained an importance which escaped previous development targets? Three reasons suggest themselves. First is the establishment of poverty at the top of the development agenda during the 1990s. During the past decade nearly all development agencies have re-affirmed and strengthened their commitment to poverty reduction and have been searching for ways in which to realise this commitment.⁶ Adoption of the IDTs is one such way, especially if the targets are internalised in the way in which DFID has done. Second, the 1990s saw an emerging emphasis on results-based management (RBM), especially in North America,⁷ with the approach being adopted by the World Bank (1993). Whilst donors have typically had some sort of monitoring at the project level (albeit often with an input-focus),⁸ the rise of RBM has shifted the focus to country programmes and the

⁵ A set of 48 indicators has been developed in support of the MDGs (DAC, 2001).

⁶ Of course poverty has always been a concern for development agencies, though the extent to which it has received explicit attention has varied greatly across time and agencies. Why poverty regained foremost position in the 1990s is beyond the scope of this paper. However, it is likely that the end of the Cold War released development aid from its political strait-jacket allowing greater autonomy for agencies.

⁷ In 1997 a special issue of the *Canadian Journal of Development Studies* was devoted to the use of results-based management in the Canadian aid agency, CIDA. The US agency USAID also has extensive experience which is drawn on later in this paper. A more comprehensive review is DAC (2000), which is also drawn on in this paper.

⁸ Project monitoring is both by the project management and, often drawing on that, by agency staff for their own purposes through supervision missions, mid-term reviews and the like. For example, the World Bank has Project Performance Reports (PPRs, formerly Form 590s) completed on an approximately annual basis for each project following a staff mission to the project. PPRs include an

agency's overall performance (DAC, 2000: 18). Results-orientation has taken hold in the UK, to some extent in the 1990s (notably the Citizen's Charter), but has been warmly embraced by the Labour government, starting with the seven pre-election pledges in 1997 and continuing with, for example, the promotion of league tables.⁹ Within government the Public Service Agreements (PSAs) are the embodiment of this approach.¹⁰ The IDTs are an obvious set of development results for development agencies to attach themselves to, and DFID has used the IDTs as the basis for its PSA targets. Finally, the IDTs have been championed from several quarters. DFID, and its Secretary of State, played an important role in this regard in establishing the targets. Once the IDTs were adopted by the World Bank (and IMF) their wide usage became assured, and they are now a reference point for the World Bank's annual flagship statistical publication *World Development Indicators*.

Whilst there is no doubt that the targets have become important, questions can be raised as to whether this is a good thing or not. This question can be asked in two ways. First, are outcome targets of any sort a good guide to policy? Second, are the IDTs a good set of performance measures? These questions are addressed in the next two sections.

3. DO TARGETS MATTER?

Targets are argued to be important for the following reasons:

- They define the objectives of policy in terms of outcomes
- They provide a basis for accountability
- They make a case for carrying out supporting activities
- They allow comparisons of performance to be made
- They create a sense of common purpose

The first point is important for two reasons. The identification of targets defines an organisation's purpose. During the past decade most development agencies have adopted an overall aim related to poverty reduction; e.g. DFID "the elimination of poverty in poorer countries" and the World Bank "a world free of poverty". Targets make a general statement of purpose more specific – in the case of the IDTs an important part of this is the recognition of the multi-dimensional nature of poverty. The focus on outcomes is also important. There has been a tendency, certainly amongst development agencies, to have performance measurement systems (monitoring and evaluation) with too strong a focus on inputs: was the money spent? Stressing outcomes reminds us what the money is being spent for, and that if the outcomes are not being achieved then having spent the money may be a bad thing rather than a good thing. This brings us to the second point of accountability.

Agencies can be held to account for their performance in achieving targets, be it firms to shareholders, government departments to tax payers or politicians to voters. The statement of a target is a commitment to achieve that target, so that the agency can be judged by whether it

assessment of performance on several criteria and overall development impact as judged against the project objectives.

⁹ See Appendix 2 of NAO (2001) for a chronology of performance measurement in UK Central Government.

¹⁰ The Public Service Agreement and Service Delivery Agreement are written commitments of a department's objectives, related performance measures, and the activities to be undertaken in support of those objectives.

does so or not. Stating targets for an organisation shows what “it is for” and so, in the case of government departments, can make a case for funding. The aid programme can be defended on the basis of the role aid monies play in reducing world poverty.

The IDTs can also be linked to the rights-based approach, popular with many development agencies including DFID. The fact that agencies and governments sign up to the targets may mean that the people whose interests they are meant to serve can lobby them to behave in the manner most consistent with meeting the targets. But it should be remembered that UN conference resolutions do not have legal status. So for a rights-based argument to be fully convincing governments need to have used its commitment to the targets to pass legislation (for example compulsory school attendance).

Where performance can be disaggregated by either service provider or different areas of provision then standardised performance targets allow comparisons to be made. Such comparisons are clearest, if debated, in the case of League Tables as used in the UK for schools. For the IDTs the relative performance of countries and regions can be compared, though it is less clear what conclusion should be drawn from this comparison. Specifically, if a country is falling short of achieving the outcome target should it receive more assistance or less, or assistance of a different kind? The answer of course depends on why the country is “off-track”, which hints at the fact that outcome-based measures do not in themselves constitute an adequate basis for performance measurement. This point forms is argued at more length later in this paper.

Finally, targets can have a motivational role both within and across organisations. DFID has IDT posters prominently displayed throughout its offices. Where several organisations are working in the same area then targets can create a sense of common purpose across these agencies. The IDTs can be argued to have played this role of the international development community, helping focus the renewed attention to poverty in the 1990s.

Targets are not without their limitations, however. Focus on a single indicator can distort programmes or sacrifice quality. For example, targets to reduce health-service waiting lists in the UK are claimed by some to have been achieved by prioritising quicker and cheaper operations. Jackson (2000) reports on a loan scheme in London where the performance target to minimise default encouraged rescheduling of ultimately unpayable debts. A target for local authorities to collect more recyclable waste was met – but many authorities then burnt the waste (NAO, 2001). The target has now been re-phrased as a target for the percentage of waste which is recycled. These concerns apply to some of the IDTs. For example, with respect to the development target for universal primary education, getting more children into already over-crowded classrooms, with few materials and poorly motivated teachers, might meet the target for quantity, but only at the expense of the quality of education provided to the children. In the case of the former target for having national strategies for sustainable development it does not capture whether the plans are being implemented or not.

Performance targets may also tend to simplify problems, focusing on what is identifiable and measurable whilst ignoring what really matters but is more complex. This is an argument which will resound with those working in poverty analysis (see, for example, Jodha, 1988, and Chambers, 1997). Even if the target is well-identified targets may have adverse organisation effects by discouraging innovation. Managers will rely on tried and tested methods rather than risk missing targets by trying out new approaches.

Targets may also undermine other forms of accountability. In the bid to satisfy performance criteria particular problems or interests may be missed. Finally, the focus on outcomes may deflect attention from the costs borne in achieving those outcomes, so that the efficiency focus of traditional management systems is lost (Cummings, 1997).

4. JUDGING THE IDTs AGAINST CRITERIA FOR PERFORMANCE MEASURES

4.1 Desirable properties of performance measures

The literature identifies several desirable features of performance measures.¹¹ The main ones are:

- *Relevant and balanced.* Balance breaks down into three areas: (i) coverage: the range of measures should cover all elements of an organisation's activities – and they should not include things which are not a part of them (i.e. indicators should be relevant), (ii) focus on both short and long-term performance (which may correspond to the distinction between outcomes and impact); and (iii) logic: measures should cover the whole process of producing the final performance measures i.e. also include inputs, activities and outputs. This last point is further discussed below. A final question of balance is the different weight given to different measures – some can be more important than others and this fact should be explicitly recognised.
- *Measures known, understood and trusted.* For performance measures to have their desired organisational impact it is necessary that they are in the first instance known by the members of the organisation and their meaning understood. But it is also necessary that they are trusted. That is, they are believed to be useful and well-defined measures and the quality of the underlying data is believed to be good.¹² A related issue is that of ownership. Ideally measures should be owned by those responsible for delivering the performance, which means they are involved in developing the targets. In practice performance measurement systems are often passed down from the top or designed by outsiders.
- *Affected and attributable.* Changes in performance measures should be affected by the activities of the organisation and the extent of the effect measurable (attributable). This is also referred to as the problem of additionality. The UK Treasury defines additionality as “the amount of output from a policy as compared with what would have occurred without intervention” (quoted in Jackson, 2000: 11) For outcome indicators attribution can be problematic and is a major issue in relation to the IDTs (see below).
- *Achievable:* targets should be achievable, but not too easily.
- *Linked to existing management systems.* Organisations will already have in place management information systems. The rise of results-based measures was in part a response to the fact that existing systems focused on inputs and internal activities, e.g. spending and staffing, rather than achieving outcomes. But new systems should not be separate or parallel to these existing systems. Rather they should be integrated, hence encompassing the logic of how inputs lead to outcomes.

4.2 Using performance-outcome measures

Performance outcome measures serve two main functions: accountability and improving organisational performance. The accountability function is straight-forward. An organisation commits itself to achieving certain outcomes and it either does or doesn't deliver. For tax-payers, politicians or shareholders this may well be sufficient and they can respond by

¹¹ This discussion draws on several sources; e.g. NAO (2001), Jackson (2000) and Hakes (2001).

¹² For a discussion of data quality issues see Divorski and Scheirer (2000).

withholding votes or funds. But managers need to understand why outcomes have or have not been achieved. There are three questions here, but only one answer. The questions are: (1) how can performance measures help change organisational practice to enhance the likelihood that targets will be met, (2) how can measures be interpreted to understand how inputs have or have not led to the desired outcomes, and (3) how can the organisation's activities be linked to the outcomes? The one answer is the importance of logic models.¹³

Ideally performance measures affect organisational culture. Simply the shift to performance-based measurement is itself a substantial change in practice: "the cultural change required to achieve performance-based management in any public or non-profit agency presents an enormous challenge" (Schierer and Newcomer, 2000: 63). But the process of planning how targets will be achieved opens up substantial possibilities for changing working practices. This planning process should be systematic and broadly-based within the organisation. This is where logic models come in.

Logic models and the logical framework are schemes for linking inputs to outcomes. Terminology varies somewhat, but all capture the same basic idea. The stages recognised in the NAO's model are resources, inputs, processes (often called activities), outputs and outcomes (also called impacts) (NAO, 2001: 2). It is a rather obvious truism that resources should be utilised in such a way as to achieve the desired outcomes. But the failure to do this has been commonly observed. For the government sector the most substantial evidence base comes from the US since the 1993 Government Performance and Results Act (GPRA) requires all government agencies to have outcome-based targets against which their performance is to be judged. The General Accounting Office (GAO) makes annual assessments of plans and outcomes. Both the GPRA itself and GAO (1998) suggest performance measures should span inputs, outputs and outcomes. However, in practice such logic models have been applied in few GPRA plans (Scheirer and Newcomer, 2000). The absence of indicators reflects the fact that there is too little conscious analysis as to how plans and programmes will affect outcomes (GAO, 1999).

The attraction of logic models is partly that they should force the agency to examine programmes to see if it really will achieve the desired outcomes (Millar *et al.*, 2000). However, logic models are a starting point for planning, rather than the end point. It is an easy matter to put a poverty-related goal at the top of a log-frame as a goal, but a rather more complex one to be clear how the inputs will help achieve that goal and to monitor progress accordingly. For example, DFID's budget support to Kenya in 2000, which was paid in support of the public sector reform programme (i.e. to finance retrenchment), included the maternal mortality rate as a goal-indicator. It is very difficult indeed to trace any links between this outcome and the activities, particularly given their short-term nature. It is all too easy to explain away discrepancies between out-turns and targets with reference to problems of attribution or external factors. But, used properly, logic models help overcome these problems.

Logic models provide a basis for understanding performance. That is, solving the problem of attribution. In principle, various types of modelling can be undertaken to demonstrate the determinants of outcomes and thus the contribution made by an organisation's activities. In practice such modelling is too costly, cumbersome and frequently contentious to provide a basis for regular performance monitoring. Hence less rigorous strategies have been proposed to establish "plausible association" between program efforts and performance (Scheirer and Newcomer, 2000: 68). These should include attempts to account for the influence of external

¹³ Logic models is the preferred term in the literature, these are very similar to the concept of the logical framework familiar in development agencies. Cummings (1997) distinguishes logic models, the logical framework and results-based management only to conclude that the three are closely related.

factors, which may of course be positive or negative. The main method here is based on having a logic model of how desired outcomes are produced.

Attribution becomes harder as we move along the causal chain. It is easy to attribute responsibility for delivering inputs, and usually for carrying out activities, although external factors may play a part. These activities should lead to desired outputs which deliver the target outcomes, again subject to external factors. If the underlying model is correct then indicators should capture if the organisation is doing what it needs to do to achieve the outcomes – which may well be the case even if the targets are not met. As a development of this approach, Scheirer (2000) argues for use of logic models to look at the relationships between measures, especially across time. Such a method is very similar to theory-based evaluation (see in particular Weiss, 1998) which is based on creating models of how outcomes are to be achieved – although there are important differences between performance measurement and the evaluation function (see DAC, 2000: 13-14).

4.3 How do the IDTs do?¹⁴

Relevance and balance

The IDTs score highly on relevance. They capture some of the main aspects of poverty in the developing world. Maxwell (1999) argues that the IDTs simplify the complex phenomenon of poverty by reducing it to a dollar a day. But this argument misses the point that every one of the targets is poverty-related, not simply that for income-poverty. Rather the IDTs should be praised for capturing the multi-dimensional nature of poverty.

It can further be argued that the IDTs have made themselves relevant. Their prominence has resulted in a consensus around this set of indicators for monitoring poverty reduction, paving the way for an unprecedented degree of co-ordination amongst donor agencies, opening the way for a harmonisation of performance monitoring which has proved difficult to achieve in the past (DAC, 2000: 21).

A more pertinent critique may be the focus on the measurable. Although the IDTs also mention the importance of qualitative aspects of development most discussion of the indicators themselves focuses on those which can be measured (and the qualitative aspects were not included in the list of the IDTs in the second White Paper). However, agencies are active in the qualitative aspects.

But the IDTs are not balanced in that for most of them the target date is 2015, with a couple for 2005. Fifteen years (twenty when they were set) is definitely a long-term goal. Although, as has been done (see the joint IMF, World Bank and UN Publication *A Better World for All*), progress can be measured with respect to being on-track, it is preferable to set explicit short-term goals. These short-term goals may either be interim targets for the same variables, or targets for outputs which will help achieve the desired outcome. This brings us to another lack of balance in the indicators.

The IDTs are also not balanced in that they are mostly outcome indicators, as are most of the wider set of indicators. Some are output indicators, but not in a consistent manner of a set of indicator capturing inputs, process, outputs and outcomes for a single target. Hence the IDTs do not provide a basis for monitoring performance on taking the steps necessary to achieve the outcomes they contain. This lack is not because such indicators are not amenable to international agreement. For example, the Copenhagen Social Summit launched the 20:20 initiative, by which 20 per cent of government spending and 20 per cent of aid monies should

¹⁴ Sections 4.3 and Part 5 draw on a questionnaire completed by a number of key informants within DFID.

be devoted to providing basic services to the poor. And the data are available to monitor progress in meeting this target (and this progress has been dismal, especially on the part of donors). Another example would be immunization, an important factor for child survival, and which has suffered set backs in recent years. A final example would be the establishment of adequate food security systems. Although nutrition is not amongst the main target indicators it is in the wider variable list and is an important determinant of some of the outcomes which are listed, most notably child survival.

Well-defined and measured

Most of the IDTs are clearly defined. Four exceptions may be noted. First, as already discussed, access to reproductive health is not measured by any existing indicator and the proxy of contraceptive prevalence is problematic. Second, the targets are loose with child survival terminology in a way which would upset demographers. Infant mortality is the probability of dying before the first birthday, child mortality between first and five birthdays and under-five mortality between birth and fifth birthday. Under-five mortality is thus a weighted average of infant and child mortality, but not liked by many demographers as the factors underlying the two mortality rates vary. The most appropriate indicators would therefore be infant and child mortality, but the targets are for infant and under-five mortality (although some sources give child mortality instead). Third, the environmental target speaks of reversing current trends in resource use. Taken literally this means that the available quantity of environmental resources should begin to increase rather than decrease. Aside from being unattainable, this target does not fit with the general consensus that it is okay to use environmental resources but in a way which is consistent with overall sustainable development. DFID, which uses the expression “managing environmental resources”, recognises this fact. Finally, it may be questioned if equality in school enrolments is an adequate proxy for gender equality. A defence of the measure is that equality of education is a necessary starting point for achieving other forms of equality.

Knowledge of the IDTs amongst DFID staff is high.¹⁵ This does not imply that they are aware of the precise technical definition of the variables. For example, understanding of the purchasing power parity (PPP) dollar a day used for income-poverty is mainly restricted to economists¹⁶ and the relevant age groups for infant and child mortality to health advisors. But it is doubtful that lack of this precise knowledge impedes work on the targets.

The chosen indicators are all ones for which data were already being collected. They are now published in a variety of places, notably the DAC website, the World Bank’s *World Development Indicators* and DFID’s *International Development Statistics*. However, data quality varies by indicator. An issue for all the indicators is coverage, meaning for how many countries data are available. The OECD document *Methodological Note* reports the baseline data for the various indicators. For some indicators coverage is low (see appendix 1 for a summary). For example, income-poverty data were missing for about 30 per cent of the population of the developing world, and nearly two-thirds of those in sub-Saharan Africa. For net primary enrolments there is only 12 per cent coverage for South Asia and 69 per cent globally.

A further issue is the frequency and timeliness with which data are available. Income poverty data come from income and expenditure surveys which are not conducted annually in many

¹⁵ This statement is based on the assessment of the DFID staff interviewed and on the author’s experience leading a poverty training programme targeted at all DFID staff.

¹⁶ The income poverty line is a dollar a day, but a dollar buys much more in, say, Kampala or Delhi than it does in New York. The poverty line is equal to the local cost of purchasing the goods which could be purchased for one dollar in the US.

developing countries but rather every 3-4 years. Vital registration systems and health facility-based reporting have inadequate coverage of the population in developing countries to be a reliable source of health data. Good quality data on child health, including infant and child mortality, are provided by Demographic Health Surveys, but these are conducted about every four years and not in all countries. It is difficult to respond to performance data which are available only with a three year lag.

Even where data are available they may be of poor quality. From the targets themselves maternal mortality data are the most problematic area. These data are notoriously unreliable (see Box 1), to the extent that is surprising the indicator was deemed suitable for inclusion amongst the targets. In measuring progress on this target *A Better World for All* reported the percentage of attended births rather than maternal mortality itself.

Box 1 Data quality: the case of maternal mortality

Maternal mortality for Ghana jumped from 400 to 1,000 from one issue of the World Bank's *World Development Report* (*WDR*) to the next. Maudlin (1994) showed that, although they both used the same source, the *WDR* reported data for 56 developing countries, whereas the UNDP's *Human Development Report* (*HDR*) did so for 55 of these 56 and 48 further countries (for which the *WDR* indicated that data were unavailable). The DAC baseline data sheets state that coverage for this indicator includes practically every country in the world.

Counting differences of less than 50 points as the same, *HDR* gave higher values than *WDR* for 26 countries. Lower for 12 and about the same for 17. Some differences are substantial, for example Benin at 800 and 161 in the two sources. The correlation coefficient between the two sets of figures is only 0.7, dropping to 0.4 for high mortality countries. A comparison of *WDR* and DHS data for 1997 also shows substantial discrepancies.

Source: Howard White *et al.* (2001) *African Poverty at the Millennium* [World Bank: Washington D.C.].

But the setting of the IDTs has in itself provided an impetus to improving data quality. As attention is focused on these outcomes there is an awareness that efforts need be made to ensure the underlying data are sound:

a more concerted effort by the donor community is needed to support partner countries' capacity to collect data and monitor progress towards the international goals over the coming years. (DAC, 2000: 22)

Are the targets achievable?

Whether or not the targets can be achieved has been the area of analysis which has attracted most attention. The issue has been looked at through the extrapolation of trends, different scenarios mainly related to economic growth, and modelling the determinants of the different target variables. The results of these analyses are summarised in appendix 2, with the main conclusions discussed here.

The main point is that there is common agreement that most the targets will not be met: "on current trends, none of the international development goals on health and education are likely to be achieved at the global level" (IMF *et al.*, 2001: 12). For example, on current trends there will still be 100 million children of school-going age out of school in 2001. These pessimistic conclusions arise even from "base run" predictions which are typically based on the usually rather optimistic growth forecasts produced by the World Bank and IMF. Forecasts based on

historical growth performance show even greater divergence from the targets, and current economic performance suggests that growth from the 1990s may over-estimate that for the new decade.

The second point is that performance varies greatly by region. The targets themselves are defined at various levels of aggregation. Some apply globally (e.g. income-poverty), some are country specific (e.g. infant and child mortality) and some necessarily apply to each country (e.g. universal primary education (UPE) will only be achieved when it is achieved in all countries). But even the targets stated as global figures are intended to be met at the disaggregated level, i.e. for each country and region: “while expressed in terms of their global impact, these goals must be pursued country by country” (DAC, 1996: 2). Here performance varies: “the education and gender equality goals are likely to be achieved in some regions and many countries. And a few countries are on track to achieve large reductions in infant and under 5 mortality” (World Bank, 2001: 12). In general, Africa is performing worst with respect to the different targets and East Asia the best. Some of these differences are explained by continued high growth in East Asia (so that the income-poverty target has already been met) compared to low growth in Africa. But African countries also suffer from much higher rates of HIV/AIDs, which has helped reverse the long-run decline in mortality rates in several countries, and from extensive conflict which undermines the ability of countries to attain any of the targets. Experience varies for the other regions. The transition economies of Eastern Europe and the former Soviet Union saw very sharp increases in poverty in the early 1990s so that the income-poverty target is unlikely to be met, and in some countries, notably Russia, social indicators have also worsened.

Whether or not the targets will be achieved is not the same thing as whether they are achievable. The papers reviewed offer limited advice on this point. The main argument made is the importance of a good policy environment (see *A Better World for All*, Demery and Walton, 1999, and Collier and Dollar, 2000): “whether or not poverty incidence will be halved by 2015 depends in part on how well economies are managed” (Demery and Walton, 1999: 83). The argument is that better policies promote growth which reduce income-poverty and other forms too such as mortality rates. Most the papers do not discuss an explicit role for development agencies. The exception is the paper of Collier and Dollar which builds on the suggestion that aid works best at promoting growth, and so reducing poverty, when the policy environment is right. Hence the prospect of meeting the IDTs is enhanced if aid is concentrated on poor countries with good policies. This piece of research has had an influence on several development agencies – including being cited in the second White Paper and being embodied in DFID’s PSA/SDA – though its theoretical and empirical basis is highly contestable (see Lensink and White, 2000).

The exception to this generally pessimistic picture is the paper of Bloom *et al.* (2000). Rather than model social outcomes as a function of income, they model income as a function of social well-being as measured by life expectancy. On the basis of the UN’s demographic forecasts they predict that the income-poverty target will easily be met.

Attribution

The IDTs do least well with respect to attribution. It is impossible (or at very best virtually impossible) for an individual agency to isolate its impact on global, or even country, trends in the IDT indicators. This fact is demonstrated by the experience of USAID.

The US Government Performance and Results Act of 1993 required government agencies to set outcome-based indicators against which their performance could be measured. In 1997 USAID laid out six strategic development goals, e.g. “broad-based economic growth and agricultural development encouraged”, and for each of these defined a set of outcome indicators at both country and global levels (e.g. “average annual growth rates in real per

capita income above 1 per cent”). With respect to the growth goal, the FY 2000 performance report states that “nearly 70 per cent of USAID-assisted countries were growing at positive rates in the second half of the 1990s, compared with 45 per cent in the early part of the decade” (USAID, 2001: v). However, that same performance report noted that “one cannot reasonably attribute overall country progress to USAID programs” (USAID, 2001: viii). Commenting on the previous year’s USAID Performance Report, the GAO had similarly observed that the goals were “so broad and progress affected by many factors other than USAID programmes, [that] the indicators cannot realistically serve as measures of the agency’s specific efforts” (GAO, 2000: 1-2). In response to these criticisms the FY 2000 performance report announced that the indicators related to the strategic goals will no longer be used to measure USAID’s performance (but they will be reported as being of interest in their own right, being referred to as “Development Performance Benchmarks”). Rather performance will be measured against the strategic objectives of the individual operating units (e.g. country programmes).

The difficulties of attribution are further illustrated by a GAO report on USAID’s child survival programme launched in the 1980s (GAO, 1996). Entitled *Contributions to Child Survival Are Significant but Challenges Remain* it lists the uses of funds under the child survival budget line and reports progress on reducing child mortality and immunisation rates in recipient countries (an activity supported by a number of donors). Pertinent observations are made on some of the activities financed by child survival programmes (a bridge in Mozambique and a water tower in Egypt) and that many countries with high mortality do not get funds, whereas those with low mortality continue to do so. But no attempt is made to link the inputs described to the outputs and outcomes reported. To do so is not an impossible task. There are models of the determinants of infant and child mortality which could be used to look at the trends in these determinants and how they have been affected by USAID’s activities. But it is a major undertaking – not one that can be accommodated within routine performance measurement. The best performance measurement can do is have a set of indicators spanning inputs to outcomes based on such an underlying model. It is for these reasons that evaluators have turned to the logical framework and approaches such as theory-based evaluation.

Integrated with existing management information systems

The DAC report on results-based management (DAC, 2000) identifies three approaches to agency-wide monitoring:

- Aggregating project and programme level outputs
- Aggregating project and programme level outcomes
- Reporting country-level trends

This classification may be simplified to “bottom up” versus “top down” systems. Bottom up systems take individual activities as the primary unit of analysis and aggregate performance across countries, sectors and the agency as a whole. Top down systems report on outcome indicators for a country (or at least sector in a country) and whole regions. Top down systems suffer from attribution problems – can a link be made between the observed outcome and the agency’s activities? If not then the information is of limited, if any, value, in guiding management systems. Bottom up systems, which are more strongly rooted in traditional management information systems, face problems of aggregation and linking to IDT-relevant outcome indicators.

All donor agencies have some sort of monitoring and evaluation system at the project and programme level, which should provide a basis for both feedback at the project level and “feed-up” to management. For our purposes here we are interested in the two questions suggested in the previous paragraph. First, are the data collected of a suitable form to be aggregated to give an overall indication of agency performance (and broken down at the country, regional and sectoral levels)? Second, if there is such an aggregation, does it yield information on outcomes in relation to the IDTs?

The World Bank’s rating system is an example of a system which does yield agency-wide results. All activities are rated under a number of criteria on a regular basis and upon completion. These ratings include an overall rating of if the activity has been satisfactory. Hence overall portfolio performance can be judged by the percentage of projects deemed satisfactory. Such analysis is published in the *Annual Review of Development Effectiveness (ARDE)* produced by the Operations Evaluation Department of the World Bank. This system does yield information of use to management, which has a systemised approach to identifying “problem areas”. In the decade from 1991 the percentage of “problem projects” rose from 11 to 20 per cent, resulting in the creation of the Portfolio Management Task Force whose report (known as the Wapenhans Report after the lead author) advocated a number of changes and contributed to the adoption of results-based management within the World Bank.¹⁷

Whilst a “satisfactory” project is defined as one which is substantially meeting its development objectives, this does not allow us to say anything about the contribution of the World Bank to meeting the IDTs. The information collected is simply not the right sort to provide that information. There is a misalignment between the data collected from the “bottom up” and the sort of outcomes being monitored in “top down” IDT-oriented systems. The same is true of all other agencies which do collect data of a form suitable for bottom up aggregation.

There are two possible responses to this problem of misalignment. The first is to say that it is inevitable. The problem of attribution is not going to be solved for routine monitoring purposes, so no attempt should be made to link agency performance as measured by bottom up systems with agency impact on the IDTs. That answer does not seem satisfactory for agencies, like DFID, which have pinned their performance to the IDT mask. So the second response to the misalignment problem is to resort to logic models. The bridge must be made between observing satisfactory activities and presumed impact on development outcomes. Whether this is feasible is an issue I return to below.

Conclusion

The IDTS are relevant as a measure of development progress, are well-defined and understood and data are mostly available if of variable quality. But the IDTs have two important short-comings, and one perhaps less important one. First the indicators themselves are output and outcome-oriented. There has not been international agreement on the underlying logic model to produce the indicators required to monitor inputs and processes necessary to achieve these outcomes. The partial exception is the consensus on the need for market-led economic growth, though whether this is the best way to achieve the targets is debatable. Second, it is not possible to attribute changes in these measures to the actions of development agencies, either individually or collectively. Both of these arguments point to the need for a more holistic performance measurement system, which integrates monitoring of inputs and process with that of outputs and outcomes. Finally, the targets seem unlikely to be attained.

¹⁷ See Carvalho and White (1996: 9-17) for further discussion.

Whilst the IDTs are mostly a satisfactory measure of development progress, this does not mean that they are suitable indicators to measure the performance of any single development agency. Indeed, a short period of reflection shows that the problem of attribution alone means that outcome indicators are by themselves unlikely ever to be suitable, especially ones of a global nature such as the IDTs. Hence some modification is necessary. To what extent has DFID's approach to the targets tackled these problems?

5. THE USE OF THE IDTs BY DFID

DFID has embraced the IDTs more strongly than any other bilateral donor. The targets have been given pride of place in both the 1997 and 2000 White Papers on International Development, they have featured strongly in the public pronouncements of the Secretary of State and have been promoted in various ways such as through posters. DFID staff view them as highly relevant to their work. But more important than these changes which appear to outsiders have been the internal efforts to accommodate the IDTs, principally the strategy papers and the PSA and SDA.

From target to strategy

At a workshop held to discuss the first White Paper many commentators argued that the Paper was full of good intentions but rather silent on how these intentions were to be fulfilled (see the papers in White, 1998). But since that time DFID has put in place a process for elaborating a strategy to this end. Central to this work have been the strategy papers: Target Strategy Papers (TSPs), Institutional Strategy Papers (ISPs) and Country Strategy Papers (CSPs) (see Box 2).

Box 2 DFID's strategy papers

Target strategy papers (TSPs) address a single development target (e.g. income poverty) or group of targets (health). Papers have also been prepared on areas of importance to poverty reduction though not explicitly included in the IDTs (water and urban poverty).

Institutional strategy papers (ISPs) have been, or are being, prepared for each of DFID's main partners amongst international organisations.

Country strategy papers outline the main challenge of poverty reduction in each of DFID's partner countries.

The intended relationship between the different documents is shown in Figure 1.¹⁸ TSPs have a central role, defining the long-term strategies to meet the development targets. These strategies should inform the content of the ISPs and CSPs. CSPs should also be informed by the relevant ISPs. Various processes have been put in place to ensure these links, starting with the collaborative way in which strategy papers were prepared. The geographical desks were charged with drawing up a response to the TSPs, outlining the implications for their work. In practice, most of the "new generation" of CSPs currently in use were prepared prior to either TSPs or ISPs being available so that the links shown in Figure 1 will not have operated in practice. But the question is the extent to which CSPs prepared since the beginning of 2001 are consistent with the TSPs (and indeed the PSA, which is discussed below).

TSPs might be faulted for a sectoral-bias in tackling problems thus missing cross-sectoral linkages. For example health matters for education and *vice versa*, or lack of close access to

¹⁸ The figure is from the DFID Poverty Guidance, *Bridging the Gap*.

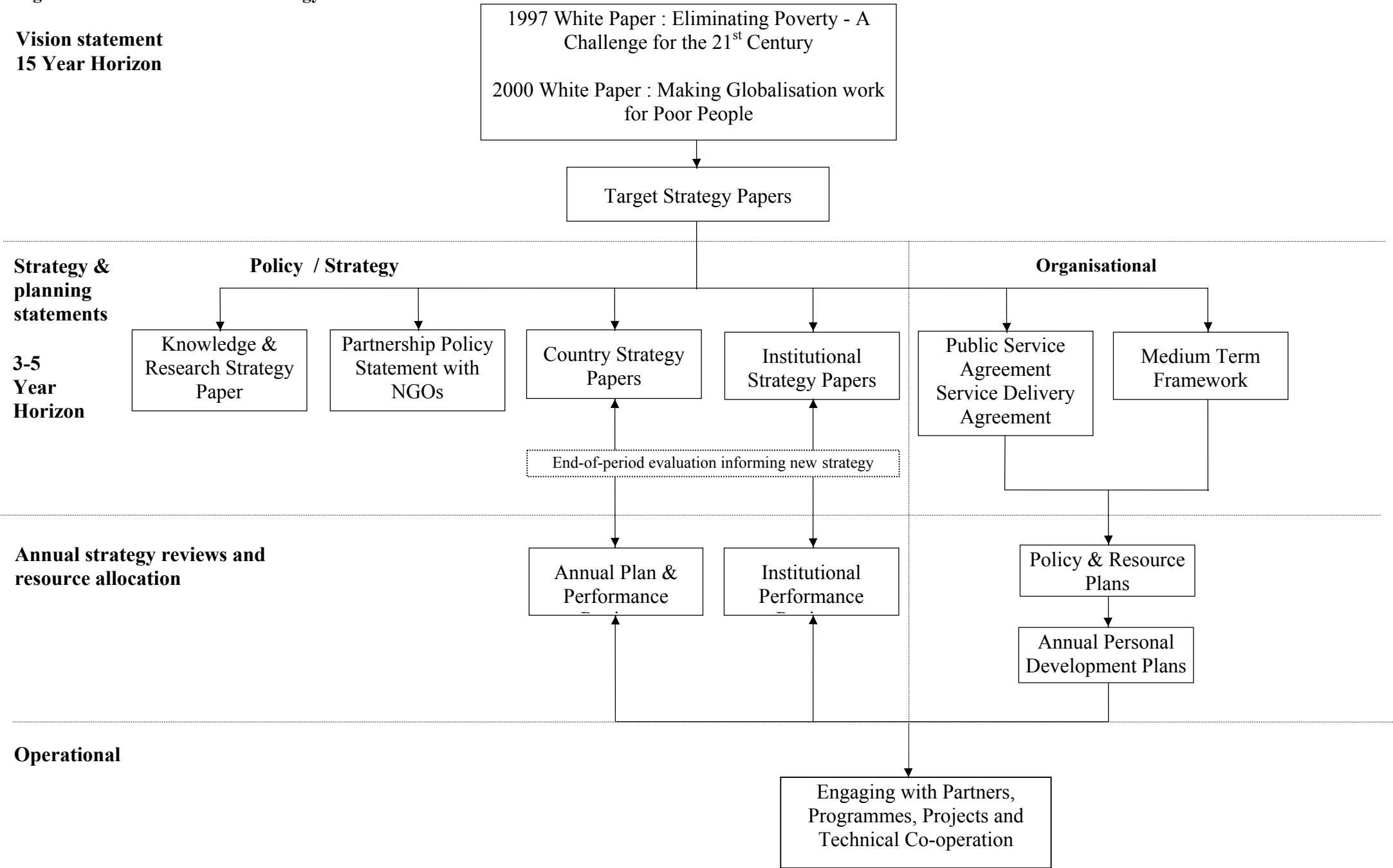
water may be the main thing keeping girls from school. Hence preparing papers on each target separately misses the importance of a multi-dimensional approach. Examining the papers shows this criticism to be only partially justified, since several papers discuss these cross-sectoral linkages. For example the environment TSP makes much of public health arguments.

CSPs might be faulted by an analogous argument. In a globalised world the country is not the appropriate unit of analysis – it should at least be the region. Whilst this argument may appear to have some merit, the fact is that DFID's operations are overwhelmingly at the country level. Hence it may it could better be argued that CSPs should have a stronger regional, or even global, dimension, and that a process should be in place for ensuring consistency of approach to the different countries in a region, rather than suggesting that regional strategy papers be prepared. Intra-regional co-ordination is largely guaranteed by DFID's organisational structure. The importance of incorporating regional dimensions may be argued to be less well catered for.

All strategy papers have a common structure which corresponds to a logic model. The first section sketches out the challenge, e.g. the relevant IDT(s) for a TSP, how an organisation relates to the poverty agenda for an ISP and the poverty situation in a country for a CSP. The next section lays out the strategy necessary to meet this challenge. This is a strategy for all partners, not DFID alone. The paper then reviews the role of different actors and, finally, the part that DFID can play.

Figure 1 Overview of DFID strategy

Vision statement
15 Year Horizon



However, an analysis of CSPs argued that they suffer from a “missing middle” (Booth and White, 1999). Whilst good on providing a poverty profile they are weaker on the causes of poverty and little attempt is made to link causes to interventions. The potential logical structure of the strategy paper is not utilised to provide the rationale for the interventions being undertaken. Although CSPs are intended to use the logical framework it does not appear in the published version, and it appears they have not been utilised very frequently in preparing the CSP itself. As argued above, the lack of a logical model undermines what scope there is for attribution.

Feeling amongst DFID staff is generally that some ISPs are weak and that CSPs are of variable quality both within specific CSPs and between countries. In addition to the problem of the missing middle outlined in the previous paragraph, CSPs often deal with the issue of partnership in a rather superficial manner. They do not always tackle head on the issue of the true poverty-orientation of either government or other partners or what will be done to influence them.¹⁹

Figure 1 shows also the PSA/SDA and the annual performance plans, which are important in bringing the IDTs down to a realistic management time-frame. Fifteen years is too long for management targets. The PSA and SDA (reproduced in integrated form in appendix 3) bring these down to targets for 2004, though sometimes making them more ambitious than their longer-run counterparts. For example the target for under-5 mortality requires an annual reduction of around 6 per cent, which over 25 years would imply a nearly 80 per cent reduction as against the IDT of 2/3. Since most countries are not on track for the longer-run target the chances of meeting the shorter-run one must be doubted. The PSA has five development objectives and one management one (value for money). Each objective has a number of performance targets, which are a mixture of process, output and outcomes. Some objectives (health and education) remain largely focused on outcome performance measures, whereas others (income-poverty) are far more process-oriented. The SDA is firmly focused on process issues linked to each of the PSA objectives.

The PSA/SDA deal to some extent with problems of affect and attribution in two ways. First, the targets are defined in relation to a smaller number of countries rather than all developing countries. For example, the health and education targets are set with respect to the top 10 recipients of UK health and education sector support respectively. Second, the PSA and SDA, at least to some extent, provide the logic model which is missing from the IDTs taken by themselves.²⁰ However, there are also problems in the approach, principally that of data availability: data of the sort required are often collected on a 3-4 year cycle at best. Hence there may well be gaps in monitoring fulfilment of many of the PSA targets. In line with the argument developed above. The outcome indicators given in the PSA are not suitable for judging agency performance. But these outcomes can be observed, and the contribution DFID may or may not have made to their fulfilment judged by the input and process indicators contained in the PSA/SDA.

The PSA shown is the second produced by DFID and differs in that the first specified targets for the top 30 recipients of the various types of aid.²¹ The shift to a smaller number reflects a more

¹⁹ Some time during 2000 the word “influence” fell into disfavour in DFID (the word is not used in the second White Paper in this context). Nonetheless “working with partners” to encourage them to DFID’s way of thinking is an important part of the new approach to poverty reduction.

²⁰ Sketching the logic models underlying the various objectives shows them to be somewhat patchy – that for health is the most complete. They are particularly weak on outputs.

²¹ Comments by NAO on an earlier version of this paper asked what was the view of UK academics of this change. I sent an e-mail question to ten academics who “do aid” (there are not so many who do). Of the

realistic objective in terms of attribution, though by no means solves that problem. On the other hand, it may be thought difficult to get representative data for a smaller number of countries. Since the target relates to specific countries proper monitoring requires data on all countries. As indicated above, these data are not likely to be readily available for several indicators. Moreover, it is not for DFID alone to develop the monitoring systems to collect these data. The IDTs potentially provide the basis for harmonised monitoring procedures across donors, although this has not been achieved in the past.

The PSA does solve the problem of how to judge agency performance by aggregating across the agency. And it probably will do so without creating burdensome reporting procedures which have weighed down other agencies:

In USAID, for example, operating units and implementing partners are beginning to complain that there is no time left for implementing programs, and that much of the higher-order results data collection is not considered directly relevant or useful to them, but is only being used to “report upward”. (DAC, 2000: 23)

But others might argue that the disjuncture between project and programme level monitoring and judging DFID’s performance is a bad move. This aggregation is not based on activity-level performance. Indeed the performance measures explicitly exclude many DFID-financed activities: for example health and education programmes not in the top ten recipients of these types of aid.

So wouldn’t the Department’s performance best be judged by aggregating the performance of different activities? In practice the data do not exist to do so, and trying to collect them would indeed be onerous. In recent years a database, PRISM, has been developed intended to contain data on all project activities. However, it is estimated that actual coverage is only about 30 per cent – and of 800 Output to Purpose Reviews (OPRs) which should be in the system only 40 actually are.²² Even if PRISM were to have full coverage it is not clear that it will be able to generate aggregate data on agency performance (in the manner of the World Bank described above), and certainly not to give any information relating to the IDTs.

At present one must wonder on what data DFID management do base their decisions. There is no “bottom up” system to indicate overall performance. And the IDT-related indicators embodied in the PSA are of little operational use.²³

However, DFID is one of the foremost donor agencies in the related developments of increased budget support and Sector Programmes (sector wide approaches, SWAPs in DFID terminology). These are consistent with both harmonised procedures and monitoring based on country-wide performance indicators. To the extent that these are developed and DFID harnesses the

seven replies, three had not heard of the PSA, one had but was unsure what it was and the other three had heard of it but were unaware of the change. DFID does not generally draw on the academic community for these “management tools” (policy documents are a different matter), and even academics with a close relationship with DFID, or a good working knowledge of aid, are usually vague on the Department’s internal workings.

²² Personal communication from Michael Flint, currently compiling DFID’s first development effectiveness report.

²³ The questionnaire sent to some DFID staff in preparing this paper asked what DFID’s response should be to the fact that *A Better World For All* showed that none of the IDTs would be met at the global level. Virtually all respondents thought no response was necessary – clearly illustrating that data on the IDTs do not yield information of operational significance.

information they provide then there will be some sort of systemised feedback, though not in a form that can be readily aggregated.

However, important issues remain as to (1) the extent to which the PSA and SDA manifest themselves in the daily work of DFID (by for example being incorporated in CSPs), and (2) if the underlying model is “right”.

The first of these questions should be a focus of NAO’s own study. DFID staff readily point to changes brought about by the renewed focus on poverty. For example, DFID’s programme to China has changed from being ATP-financed infrastructure in the relatively affluent seaboard provinces to social sectors in poorer Western provinces. It is debatable how much the change comes from adopting the IDTs *per se* rather than an increased poverty focus. Such a focus was already there since 1990. And, for example, during the nineties the Zambia programme shifted from secondary schools and support to hospitals to primary education and health clinics. Attitudes as to how important the IDTs themselves are to arguing for the poverty agenda varies between staff – but their effect is seen as either neutral or positive, there was no suggestion that they have detracted attention from “main issues”. Those that see a positive role argue that stressing the IDTs has enabled internal policy changes (abolishing ATP, more aid for South Asia) and can be a useful tool in discussions with partners.

It is striking that DFID staff and outside experts stress that a major part of DFID’s contribution to achieving the targets comes through its influence on partner country policies and the actions of other actors. Yet measuring impact through influence is an under-researched area.

Whether the model is right is not a matter of objective fact. Targets do not in themselves contain the strategy as to how they should be attained so that competing strategies may be proposed. The changes DFID is making, e.g. toward selectivity, are in line with the consensus amongst donor agencies, if disputed by some critics (Lensink and White, 2000). Other areas of contention include the nature and depth of debt relief and cost recovery schemes for basic services.

Data quality

DFID staff are aware of data quality issues. Whilst operational staff express appropriate scepticism as to the quality of the data, various initiatives are underway to promote use of the data and improve them. Indeed, an advantage of the IDTs is said to have been to draw attention to data quality issues. The publication formerly called *British Aid Statistics* has become *International Development Statistics* and includes data on the IDTs. CSPs are required to report a country’s performance with respect to the IDT indicators (though this has been done in an uneven way, see Booth and White, 1999). DFID statistical staff are well-informed on data quality and are active in supporting initiatives for them to be improved, notably DAC’s PARIS (Partnerships In Statistics for development in the 21st Century).²⁴

DFID staff don’t feel that the IDTs stress quantitative at the expense of qualitative aspects of development since they believe the importance of the latter is well understood. In addition DFID supported work by DAC on governance indicators to quantify the governance target. When the working group was unable to reach agreement DFID support has shifted to on-going efforts to develop acceptable indicators by the World Bank.²⁵ The difficulty in selecting measures of

²⁴ See www.paris.org

²⁵ For a summary of this work see www1.worldbank.org/publicsector/indicators.htm. This site includes a discussion of available indicators including those from other sources such as Freedom House.

“governance” revolves in part around disagreements as to what constitutes “good governance”. With respect to human rights European countries may wish to include the absence of the death penalty, but this is still applied in other developed countries. The extent of restrictions on individual freedom for reasons of national security is another murky area. The indicators being developed by the World Bank attempt to side step these issues to some extent to proposing a set of process indicators – such as civil service wages and the nature of elections (e.g. Proportional Representation or not) – which have no normative content. Performance measures on the quality of governance – e.g. corruption and predictability of policy-making – are included separately.

Conclusion

The PSA and SDA have potentially provided a vehicle to enable DFID to travel the road travelled in rather more time by USAID. That is, there is a move away from the outcome indicators as the measure of performance. DFID staff are indeed generally, although not universally, sceptical that the IDTS can be used to assess DFID’s performance. The PSA and SDA spell out a number of process indicators which may be used to capture DFID’s contribution toward meeting the IDTs. This role is appreciated by DFID staff. One commented that the IDTs could not be used to assess DFID’s performance but that it should “track changes and make sure that their work is consistent with the IDTs” (adding that the PSA played this role). This scepticism also appeared in the fact that no respondents thought that the fact the indicators are currently not on track to meet the targets called for major changes in DFID’s strategy. It is also shown by the ease with which the DFID Departmental Review explains away deviations from target. However, the PSA also does contain outcome indicators. So it is worth emphasising that development outcome indicators of this sort are not a suitable vehicle for judging the performance of individual agencies. And there are problems of incomplete coverage of DFID’s activities and misalignment with the reporting system being developed using PRISM.

This paper has not been able to assess the extent to which the model contained in the PSA/SDA has influenced DFID’s work on the ground.

6. Conclusions

The International Development Targets, launched in 1996, have caught the attention of the development community. Unlike previous targets, they have not fallen by the wayside, but continue to be referred to and monitored. Their dominance is explained partly by the renewed focus on poverty, partly by the rise of results-based management and partly by the support they have received from key actors including the UK government. Targets can play an important role in accountability and performance measurement, though they are not without their disadvantages. To be most effective, performance measures should satisfy a number of criteria. The IDTs satisfy only some of these criteria. They are very relevant, mostly well-defined and correspond to existing indicators. But they are mostly outcome-oriented with little effort made to build a consensus around an underlying logic model of how the targets are to be achieved. In the absence of such a model it is extremely difficult to say anything sensible to attribute changes in target indicators to the actions of the development community. To put it bluntly: the IDTs are not suitable for judging the performance on individual development agencies.

Amongst bilateral donors, DFID has been a prominent supporter of the targets and has made substantial steps to internalise them. The targets have been central to the two White Papers produced since 1997 and strategy papers have been produced as to how the targets may be achieved. The PSA and SDA contain performance measures which are related to the IDTs, but

also include inputs, process and outputs. Hence they contain a model of how to achieve these interim targets and so, implicitly, the IDTs themselves. To the extent that performance is judged by these intermediary indicators, then the PSA/SDA represent an improvement. However, worries remain over the misalignment between these top down targets (which have incomplete coverage of DFID's activities) and the nascent bottom up system in PRISM. A key area of further investigation is the extent to which the SDA affects the work of the various parts of DFID.

References

Baulch, Bob (1996) "The new poverty agenda – a disputed consensus" *IDS Bulletin* **27**(1): 1-10.

Bloom, David, David Canning, Bryan Graham and Jaypee Sevilla (2000) "Out of poverty: on the feasibility of halving poverty by 2015" CAER Discussion Paper No. 52

Carvalho, Soniya and Howard White (1994) "Indicators for Poverty Reduction" *World Bank Discussion Paper* [Washington D.C.: World Bank].

Carvalho, Soniya and Howard White (1996) "Implementing Projects for the Poor: what has been learned?" *Directions in Development* [Washington D.C.: World Bank].

Chambers, Robert (1997) *Putting the Last First* [London: IT Publications]

Collier, Paul and David Dollar (2000) "Can the World Cut Poverty in Half? How Policy reform and effective aid can meet the IDGS"

Cummings, F. Harry (1997) "Logic Models, Logical Frameworks and Results-Based Management: contrasts and comparisons" *Canadian Journal of Development Studies* **18** 587-596.

DAC (1996) *Shaping the Twenty-first Century* [Paris: OECD].

DAC (2000) "Results Based Management in the Development Co-Operation Agencies: a review of experience (Executive Summary)" [Paris: OECD]

DAC (2001) "Millennium Development Goals" [Paris: OECD].

David, Booth and Howard White (1999) "How Can Country Strategies Be A More Effective Tool for Poverty Reduction?: Issues From A Review Of The "New" CSPs", mimeo.

Divorski, S. and M.A. Scheirer (2000) "Improving data quality for performance measurement: results from a GAO study of verification and validation" *Evaluation and Program Planning* **24**(1) 83-94.

Demery, Lionel and Michael Walton (1999) "Are the poverty and social goals for the 21st century attainable?" *IDS Bulletin* **30**(2) 75-

General Accounting Office (1996) *Contributions to Child Survival are Significant but Challenges Remain* [Washington D.C.: GAO].

General Accounting Office (2000) "Observations on the US Agency for International Development's Fiscal Year 1999 Performance Report and Fiscal Years 2000 and 2001 Performance Plans" [Washington D.C.: GAO].

Hakes, J.E. (2000) "Can measuring results produce results: one manager's view" *Evaluation and Program Planning* **24** 319-327.

Hanmer, Lucia and Felix Naschold (2001) "Attaining the IDTs: will growth be enough?", mimeo, London: ODI.

IMF, OECD, UN and World Bank (2000) *A Better World For All*, Washington D.C.

Jackson, Annabel (2000) "An evaluation of evaluation: problems with performance measurement in small business loans and grant schemes" *Progress in Planning* **55** 1-55.

Jodha, N.S. (1988) "Poverty debate in India: a minority view" *Economic and Political Weekly* **22**(45-47): 2421-28.

Lensink, Robert and Howard White (2000) "Assessing Aid: a manifesto for aid in the 21st century?" (with Robert Lensink) *Oxford Development Studies* 28(1), 2000.

Millar, Annie, Ronalf Simeone, and John Carnevale (2000) "Logic models: a systems tool for performance management" *Evaluation and Program Planning* **24**(1) 73-81.

National Audit Office (2001) *Measuring the Performance of Government Departments* [London: The Stationery Office].

Scheirer, Mary Ann (2000) "Getting More 'Bang' for Your Performance Measures 'Bucks'" *American Journal of Evaluation* **21**(2) 139-149.

Scheirer, Mary Ann and K. Newcomer (2000) "Opportunities for program evaluators to facilitate performance-based management" *Evaluation and Program Planning* **24**(1) 63-71.

Ulrich, Dave, Jack Zenger and Norm Smallwood (1999) *Results-Based Leadership* [Boston: Harvard Business School Press].

UN (2000) *United Nations Millennium Declaration* [New York: UN] <http://www.un.org/millennium/declaration/ares552e.pdf>

USAID (2000) *FY 2000 Performance Review* [Washington D.C.: USAID].

Weiss, Carol (1998) *Evaluation* [Upper Saddle River, N.J. : Prentice Hall]

White, Howard (ed.) (1998) Special issue of *Journal International Development* on 1997 UK White Paper on International Development.

White, Howard (1999) "Global poverty reduction: are we heading in the right direction?", *Journal of International Development* **11** June, 1999.

World Bank (1993) *Getting Results: the World Bank's agenda for improved development effectiveness* [Washington D.C.: World Bank].

World Bank (2001) *Poverty Trends and Voices of the poor* [Washington D.C.: World Bank]
<http://www.worldbank.org/poverty/data/trends/index.htm#intro>

Appendix 1 Data coverage of baseline data for the IDTs

		East Asia and Pacific	Europe and Central Asia	Latin America and Caribbean	Middle East and North Africa	South Asia	Sub-Saharan Africa	World
Number of countries		20	27	35	15	8	48	200
Poverty Headcount	Countries	6	9	9	5	4	11	44
	% of pop	83	71	65	52	78	36	71
Poverty depth	Countries	4	16	10	3	4	11	48
	% of pop	86	71	65	52	78	36	57
Poor's consumption share	Countries	8	16	15	7	5	18	69
	% of pop	86	72	73	57	86	58	86
Underweight	Countries	12	10	25	1	8	45	120
	% of pop	98	66	99	28	100	100	89
Net primary enrolment	Countries	9	16	24	13	2	30	127
	% of pop	88	73	90	97	12	56	69
Completion 4th grade	Countries	11	8	17	9	4	32	100
	% of pop	84	28	81	50	78	61	68
Literacy	Countries	8	12	25	13	7	36	120
	% of pop	97	68	100	89	100	83	83
Girls to boys enrolments	Countries	14	25	25	13	7	42	161
	% of pop	94	99	66	97	100	98	96
Female to male literacy	Countries	28	8	6	10	21	4	77
	% of pop	78	95	94	63	98	3	84
Infant mortality rate	Countries	20	27	35	15	8	48	200
	% of pop	100	100	100	100	100	100	100
Under five mortality rate	Countries	20	27	33	15	8	48	187
	% of pop	100	100	99	100	100	100	100
Maternal mortality rate	Countries	12	26	23	11	5	38	144
	% of pop	100	99	99	97	98	97	100
Attended births	Countries	14	18	14	12	8	39	137

		East Asia and Pacific	Europe and Latin America Central Asia and Caribbean	Middle East and North Africa	South Asia	Sub-Saharan Africa	World	
Contraceptive prevalence	% of pop	42	99	62	97	100	88	86
	Countries	5	6	15	8	3	29	69
Process for sustainable development	% of pop	90	43	61	76	95	84	75
	Countries	34	7	16	5	18	6	86
Access to safe water	Countries	21	9	33	13	7	45	151
	% of pop	100	26	100	85	100	89	89
Land area protected	Countries	27	12	23	13	5	39	146
	% of pop	100	100	99	100	98	98	100
GDP per unit of energy use	Countries	6	24	21	12	5	18	116
	% of pop	95	97	96	100	98	76	95
Carbon dioxide emissions	Countries	20	26	35	13	9	45	..
	% of pop	100	100	100	100	100	99	..
Source: DAC Methodological Note DCD/DAC(98)6/ADD								

Appendix 2 Prospects of meeting the IDTs

	Current trends (<i>Better World for All</i>)	Trends and scenarios (World Bank, 2001)	Modelling (Demery and Walton)
Half income poverty	On track overall, but progress in Asia not in Africa, and inequality a barrier in Latin America.	Met in 'base case' overall, though not for Africa. In 'low case' (actual growth in 1990s), not met overall, though is achieved in East Asia.	Achieved in only half of the 36 countries studied (but representing 86% population) with unchanged policies but in 28 with improved policies.
Universal primary education	Enrolment rising but not fast enough. On current trends 100 million school age children will be out of school in 2015.	Not likely to be achieved, more than 100 million school age children will not be in school in 2015. In Africa the <i>percentage</i> not in school will increase.	..
Gender equality in enrolments	Gender gap narrowing more slowly than required to meet target.	In 1995 girls accounted for 43 per cent of primary enrolment and 40 per cent of secondary. On current trends both these figures will be 47 per cent by 2005.	..
Two-thirds reduction in infant and child mortality.	Only one in ten countries achieving sufficiently rapid progress to meet the target.	Of the 151 countries with adequate data, only 26 lowered infant mortality fast enough in the 1990s to reach the target, and 11 experienced increased mortality. For under 5 mortality these last two figures are 17 and 14 respectively.	Based on expected trends in female education and income under five mortality will fall but remain, on average, 60 per cent higher than the 2015 goal.
Three-quarters reduction in maternal mortality	Proportion of attended births rose only slowly in the 1990s, far less than required to meet target.	There has not been a significant improvement in maternal mortality in the 1990s.	..
Reproductive health	No quantified target, but increased access indicated by increased contraceptive use.
Environmental sustainability	Less than half the world's countries have adopted strategies and fewer are implementing them.

Appendix 2 Prospects of meeting the IDTs ctd.

	Modelling (Hanmer and Naschold)	Modelling (Collier and Dollar)	Modelling (Bloom <i>et al.</i>).
Half income poverty	Target met in East Asia under all scenarios, for other regions only under more optimistic assumptions and very difficult for Africa.	On current trends overall target will be achieved, but concentrated in East and South Asia. But with policy reform and more efficient aid target could be met in all regions	Based on link from health (life expectancy) to poverty the target can be met by relying on current trends.
Universal primary education	Growth-based forecasts suggest will not be achieved in South Asia, Africa or Middle East.
Gender equality in enrolments
Two-thirds reduction in infant and child mortality.	Infant mortality reduction achieved overall on optimistic assumptions about better health, though not in South Africa. If HIV/AIDS spreads then only met in East Asia, with a 24 per cent decline overall. Similar picture for under 5 mortality, though target just missed even under best case scenario.
Three-quarters reduction in maternal mortality	Not met in any of the cases, with a 56 per cent decline in the best case scenario.
Reproductive health
Environmental sustainability

Appendix 3 DFID Public Service Agreement and Service Delivery Agreement

Aim: The elimination of poverty in poorer countries

Objectives	Performance targets	Delivery
<p>I To reduce poverty through the provision of more focused and co-ordinated development assistance by the international community to low and middle income countries.</p>	<p>1 An increased focus by DFID on poor countries, particularly those with effective governments pursuing high growth and pro-poor economic and social policies, as demonstrated by:</p> <ul style="list-style-type: none"> • an increase in the percentage of DFID's bilateral programme going to poor countries, particularly those with favourable policy environments; • an increase in the percentage of EC development assistance going to poor countries and; • adoption and implementation of effective Poverty Reduction Strategies by 2004 in all countries accessing International Development Agency (IDA) high impact or adjustment lending. 	<p>Successful delivery depends on DFID and multilateral institutions becoming more selective, and focused on poverty reduction. DFID will therefore:</p> <ul style="list-style-type: none"> • Deliver a more effective and focused bilateral programme by: (a) working with partners in poor countries to deliver country strategies which support poverty reduction, and;(b) allocating DFID support taking account of numbers of poor people, the effectiveness of country programmes, and partners' progress in developing and implementing sound pro-poor policies. • Seek to improve the effectiveness of EC development assistance and the European Development Fund by working with other Government Departments (especially FCO and HMT) and EU Member States to: (a) establish better organisation of EC programme delivery, by end- 2001; (b) gain agreement in Council and Commission to re-direct allocations and spend towards programmes which reduce poverty by 2003; (c) increase the proportion of EC country specific oda going to poor countries from 50% in 1998 to 70% in 2006. <p>Where poor countries have demonstrated a clear commitment to developing and implementing comprehensive Poverty Reduction Strategies, donors should respond by supporting delivery of these Strategies. DFID will therefore:</p> <ul style="list-style-type: none"> • Provide support to at least 12 partner countries by 2004 to develop and implement Poverty Reduction Strategies in co-ordination with other donors.
<p>II To promote sustainable development through co-ordinated UK and International Action.</p>	<p>2 To promote the integration of developing countries into the global economy through co-ordinated UK and international action, including by:</p> <ul style="list-style-type: none"> • relief of unsustainable debt by 2004 for all heavily-indebted poor countries (HIPC) committed to poverty reduction, building on the internationally agreed target that three-quarters of eligible HIPCs reach decision point by end 2000 (Joint Target With HM Treasury) and; • gaining international agreement 	<p>Relief of unsustainable debt burdens is essential if poor countries are to harness the resources they need for economic growth and development, to reduce poverty and reap the benefits of globalisation. DFID will therefore:</p> <ul style="list-style-type: none"> • Work to secure faster, wider and deeper debt relief for the poorest countries, through effective implementation of the Heavily Indebted Poor Countries (HIPC) Initiative. Effectiveness will be determined by the involvement of all creditors (including securing the necessary financing for multilateral creditors), the extent of front-loading i.e. that more of the benefit of debt relief is felt in the early years (including from Decision Point), the speed of the process and the strength of the link to poverty reduction, ensuring that debt relief assists countries to implement their national poverty reduction strategies and achieve the international development targets. <p>To ensure development is sustainable over the long term and benefits future as well as current generations, poor countries need to integrate sustainable development into their policies and programmes. DFID will therefore work towards:</p> <ul style="list-style-type: none"> • Developing guidance on the principles of sustainable development, securing OECD Development

Objectives		Performance targets	Delivery
		on the integration of social, economic and environmental aspects of sustainable development into poverty reduction programmes.	<p>Assistance Committee (DAC) agreement to it by mid-2001, and work to secure wider international agreement by end 2001;</p> <ul style="list-style-type: none"> Successful integration of these principles into government, multilateral and DFID policies and programmes in 10 key DFID partner countries by early-2004, including agreed approaches to water resources management, and capacity building for environmental management; <p>Successful integration of poor countries into the global economy will depend on the creation of a supportive environment in which trade and enterprise can flourish, and contribute to poverty reduction. DFID will therefore:</p> <ul style="list-style-type: none"> Work with the public and private sectors to improve the business environment, especially access to finance and other business services for enterprises that employ or benefit the poor. Promote increased private sector foreign investment in poor countries by turning CDC into a Public-Private Partnership, when business conditions are right, with majority private capital. CDC is required to make 70% of its new investments in poor developing countries and seeks to make 50% of its new investments in sub-Saharan Africa and South Asia. Work with the EU and other partners for multilateral trade negotiations to improve trading opportunities for poor countries, whilst working with other donors to deliver more effective capacity-building support for poor countries so that they are equipped to participate fully in the international trading system. <p>Effective action to tackle HIV is essential if poor countries are to sustain economic growth, development and poverty reduction. DFID will therefore:</p> <ul style="list-style-type: none"> Work with partners in countries with high, or increasing, HIV prevalence to develop and implement strategies which intensify multi-sector and co-ordinated international action on HIV/AIDS. <p>Successful delivery depends on governments, donors, international bodies, civil society groups, the private sector and others co-operating closely in the design and delivery of coherent, complementary policies and interventions in order to defuse tensions, reduce violence, tackle the factors that underlie armed conflict and build governments and institutions capable of sustaining peaceful and democratic societies.</p> <p>Where the UK can make a significant contribution, DFID, FCO and MOD will work in partnership with others to:</p> <ul style="list-style-type: none"> Strengthen international and regional systems and capacity for conflict prevention, early warning, crisis management, conflict resolution/peace making, peace keeping and peace building. Contribute to global and regional conflict prevention initiatives, such as curbing the proliferation of small arms and the diversion of resources to finance conflict. Promote initiatives in selected countries, including indigenous capacity building, to help avert conflict, reduce violence and build sustainable security and peace
		3 Improved effectiveness of the UK contribution to conflict prevention and management, as demonstrated by a reduction in the number of people whose lives are affected by violent conflict and by a reduction in potential sources of future conflict, where the UK can make a significant contribution (Joint Target With FCO and MoD).	
III	Improved education	4 Improved education systems in our top ten recipients of DFID education	Successful delivery depends on donors and poor country partners working together to design and delivery effective policies and support for education. DFID will therefore work in partnership with others to

Objectives		Performance targets	Delivery
outcomes in key countries receiving DFID education support		support demonstrated by: <ul style="list-style-type: none"> • an average increase in primary school enrolment from a baseline established in 2000 of 75% to 81% on the basis of data available in 2004; and • improvements in gender equality in education, particularly primary education. 	support: <ul style="list-style-type: none"> • Implementation of the agenda agreed by the International Community at the Dakar World Education Forum in April 2000 through the provision of focused support by relevant multilaterals, partner countries, bilateral donors and NGOs; • Successful adoption and implementation of education sector strategies which include explicit objectives on equitable access for girls and boys by 2004, in at least 8 of our top 10 recipients of bilateral education assistance; • Development of basic monitoring and evaluation mechanisms and their integration into education sector strategies by 2004 in at least 8 of our top 10 recipients of bilateral education assistance.
IV Improvements in health outcomes in key countries receiving DFID health care assistance	5	Improvements in child, maternal and reproductive health in our top ten recipients of DFID health care assistance demonstrated by: <ul style="list-style-type: none"> • a decrease in the average under-5 mortality rate from 132 per 1,000 live births in 1997 to 103 on the basis of data available in 2004; • an increase in the proportion of births assisted by skilled attendants from a baseline established in 2000 of 43% to 50% on the basis of data available in 2004; and • improved access to reproductive health care. 	Successful delivery depends on donors, and poor country partners working together to design and deliver effective policies and support for health. DFID will therefore work in partnership with others to support: <ul style="list-style-type: none"> • Development and implementation of strategies focused on improving access to safe water and sanitation and reducing levels of child mortality, in at least 8 of the top 10 recipients of bilateral health assistance by 2004; • Development and implementation of health sector strategies by 2004 in at least 8 of the top 10 recipients of bilateral health assistance which: (a) aim to improve child health outcomes and include actions to strengthen immunisation and prevention, and the treatment of childhood illnesses, including malaria where endemic; and (b) include explicit policy and operational frameworks to strengthen the capacity of health systems, improve the quality and coverage of maternal health care, and ensure universal access to reproductive health services; • Strengthened multilateral initiatives to combat HIV/AIDS in Africa (UNAIDS) and Roll Back Malaria (WHO) demonstrated through national strategies, with jointly agreed milestones, in at least 5 of the top 10 recipients of DFID healthcare assistance.
Value for money	6	Improved value for money and effectiveness of projects in DFID's bilateral programme, as demonstrated by a year on year improvement in the index of their evaluated success	Successful delivery of improved value for money will be measured by the index of evaluated success. This depends on annual project scoring and risk labelling of projects; roll-out and full use of Performance Reporting Information System for Management by 2001; effective quality control and monitoring.