DFID Evidence Survey

Results Report

November 2013

Produced by:

The Evidence into Action Team

Acknowledgements

The 2013 evidence survey is an initiative that is led by the Evidence into Action team but has been made possible only by a collaboration of staff from across DFID. The following people have been instrumental in enabling the success of this.

Max Gasteen, Ben Cattermoul, Kirsty Newman, Jessica Prout, Sophie Boldon, Ed Barney, Harriet MacDonald Walker, Laura Rivkin, Esther Dobson (Evidence into Action), Matthew Johnson-Edan, James Owen (Office of the Chief Economist); Kerstin Hinds, Ivan Parks (DFID Kenya); and Paul Atherton (Human Development Department).

DFID Evidence Survey 2013: Headline Findings

446 staff participated in a survey in July 2013 to explore the attitudes to, and use of evidence, in DFID. With a response rate of 81% from the 552 staff who were invited to take part, the results are broadly representative for the organisation.

- **Staff value evidence**, with 87% of staff agreeing that it is important for their work**.**

- There is broad agreement across DFID that the **use of evidence has increased over the past three years**, with staff in RED and the SCS reporting the highest increase.

- **Nearly half of DFID staff believe that spending more time on evidence has had a positive impact on DFID's work** to reduce global poverty. However, DFID needs to be more rigorous in the way that it measures and demonstrates this value.

- **Staff do not feel that the strength of their commitment to evidence is matched by an equal commitment within DFID**. Whilst 63% of staff said that they valued evidence a lot, only 40% of staff felt that DFID valued evidence to the same degree.

- **58% of staff felt that their manager supported their use of evidence**. Focus groups stated that the focus is often confined to the business case and does not extend across the programme and policy cycle.

- The **biggest barriers to using evidence are easily finding it and having enough time** to consider it. There is a strong demand from staff to increase the accessibility of evidence.

- Two thirds of staff have **confidence in their skills to find and use evidence.** But staff also want **more training on a range of subjects that is better targeted and tailored** to different needs and accessible for country offices.

- **Cadres and informal networks are a great source of knowledge and experience in DFID**. But staff would welcome a more coordinated approach to information sharing.

- DFID **services which synthesise evidence are valued** but their profile needs to be raised so that staff know they are available.

- Overall, staff who have engaged with centrally commissioned research are at least 'somewhat satisfied'. But staff raised concerns that the **research is not regionally relevant and more could be done to engage country offices and policy teams**.

- Staff recognise **that evaluation is important but want more support and for findings from evaluations to be shared**. Staff also said that DFID was 'average' at learning and applying lessons.

Contents

1. Introduction

The DFID evidence survey was set up in July 2013 to address two important questions:

i. What are the strengths and weakness of DFID as an evidence-informed organisation?
ii. What can we do to improve?

Why is this important?

DFID's vision is to become world-class in using evidence to drive value for money and impact, and to influence other partners to do the same. In October 2012, the Secretary of State set out the challenge:

> *"…I want to make sure that we invest in what works. Where we don't know, I want to find out…. I want to champion research and evidence to make sure that we are learning all we can, and we know that our approach is actually going to work…"*[1]

However, we know that we have work to do to fulfil this ambition. In 2012, the Quality Assurance Unit found that "…in 17 of 29 (59%) business cases assessed, the choice, function, judgement of quality, or use of evidence was considered lacking, inappropriate or poorly employed, requiring minor or major revisions..."[2] These findings are supported by the National Audit Office (NAO) review of business cases, which reported weaknesses in the appraisal case due to "…the requirement for good quality quantitative data, which DFID often struggles to gather and utilise. Analysis of a number of options, including 'do-nothing', in a robust quantitative and qualitative way is still a relatively new requirement for DFID and there are improvements to be made in this area…"[3]

These problems are compounded by the fact that we don't yet know enough about the major constraints people face in finding and using evidence effectively, and therefore what we can do about them. This means that our current resources, such as training and guidance notes, are not targeted as effectively as they could be.

What is the focus of this report?

This report provides an overview of the results from the 2013 evidence survey. Aimed at DFID staff, it is the first of a number of products that will support the uptake and use of the survey results.

This report sets out:

1. An **overview of the survey design,** setting out the conceptual framework that supports the survey design and implementation as well as details of the survey sample.

2. The **survey results**, focusing on:

---

[1] Secretary of State Justine Greening's speech to Conservative Party Conference, 9th October 2012
[2] Quality Assurance Unit report to the Investment Committee 2012 para 3.3
[3] National Audit Office report on Business Cases in DFID 2011 para 9.

i. Evidence value and use
ii. Management support for the use of evidence
iii. Evidence skills and knowledge
iv. Evidence availability and accessibility
v. Research design

3. **What happens next**? How the findings from this survey will be used to improve how DFID engages with, and uses evidence for, its policy and programme decision-making.

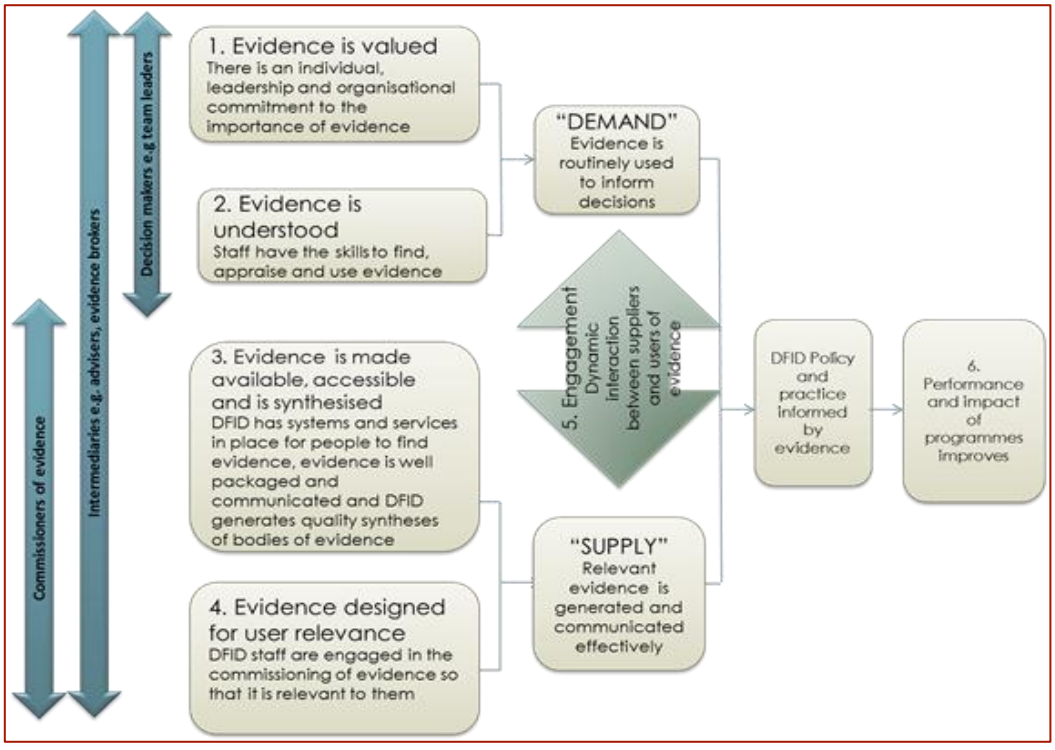
2. Evidence Survey Design

For DFID to be an evidence-informed organisation, a balance is needed between evidence demand and supply. On the demand side, evidence must be valued and staff must have the skills to access and use it. On the supply side, there are important factors such as the relevance and accessibility of different types of research that influence whether, and how, people will use it.

The conceptual framework below (*figure 1*) demonstrates how these elements come together in a dynamic relationship of evidence supply and demand. For DFID, if this relationship is working well then this leads to improved decision making and ultimately poverty reduction. In the context of the evidence survey, the statements in boxes 1-6 provided the basis for the survey design.

Like all conceptual frameworks, the diagram involves a degree of simplification. It is hard to capture coherently the overlap between producers and users of evidence and decision-makers. Advisors, for example, may often find themselves playing multiple roles that span both supply and demand functions. There are also a very complex set of assumptions that underpin the assertion that improved use of evidence will lead to improved decision making and improved programme outcomes. We hope that this survey can increase our understanding of this dynamic within DFID.

Figure 1: Evidence Supply and Demand



2.1 Survey Methods

The research process consisted of two parts – the electronic survey and a series of focus groups. Each is described below.

The Electronic Survey

The survey (*annex 1*) was structured around the six numbered statements of the conceptual framework. It included comprised of a mix of questions, including:

- ***Perception questions:*** measured on a Likert Scale of 1-5 and designed to explore variations between groups in terms of incentives, demand, skill, use and impact of evidence.
- ***Skills and knowledge questions:*** designed to test people's knowledge of different dimensions of using evidence. These questions are based on the cross-competency frameworks and use economics, statistics, research methods and evaluation skills.
- ***Space for qualitative response:*** designed to give respondents an opportunity to elaborate on their answers.

A number of the questions included in the survey were purposively similar to those included in a survey of use of evidence in the World Bank (WB). This survey was analysed and presented in 'Knowledgeable Bankers' (Ravallion, 2011). However, a key difference is that the World Bank survey was primarily concerned with use of evidence generated by internal researchers, whereas this survey was concerned with the use of both internal and external sources.

The population of the survey was all DFID staff from grade B1(D) and above, across 22 Whitehall, Abercrombie House, and Country Offices. A stratified random sample of this was undertaken based on cadre and grade.

There were 552 recipients of the survey and 446 full responses, giving an overall response rate of 81%.

Table 1: Response Numbers

DFID DIVISION	No.	PROFESSIONAL CADRE	No.
Asia, Caribbean and Overseas Territories Division (ASCOT)	62	Climate	27
Communications	5	Communications	5
East and Central Africa	60	Conflict	17
Finance and Corporate Performance Division (FCPD)	23	Economics	34
HR Security and Facilities	10	Education	28
International Finance	9	Evaluation	23
International Relations	9	Generalist	87
Joint Trade and Policy Unit	2	Governance	27
Policy Division	61	Health	29
Research and Evidence Division	48	Humanitarian	13
Top Management Group	9	Infrastructure	23
West and Southern Africa	56	Livelihoods	26

Western Asia and Stabilisation	68	Private sector	24
Other	24	Programme management	6
-	**-**	Social development	27
-	-	Statistics	31
-	-	Other	19
Total	**446**	**Total**	**446**
LOCATION OF POST	No.	**GRADE**	No.
Country office	200	A1	119
East Kilbride	79	A2	205
London	163	A2(L)	64
EU	2	B1(D)	27
Other	2	SCS	25
-	**-**	Other	6
Total	**446**	**Total**	**446**

Focus Groups

Seven focus groups were held (*table 2*). The discussion framework for the focus groups was designed around an initial overview of the electronic survey responses, enabling the survey team to drill down on the contextual aspects that had influenced the survey responses.

Table 2: Focus Group locations

Focus Group (FG)	Code
Whitehall 1	WH1
Whitehall 2	WH2
Whitehall 3	WH3
Kenya	K
Abercrombie House	AH
Democratic Republic Congo	DRC
India	I

The original participant list for the focus groups was randomly selected to enable a representative selection of staff from across DFID. However, it should be acknowledged that the challenges of encouraging people to participate in meetings held during the summer months inevitably led to a degree of self-selection bias based on interest and enthusiasm to engage in a discussion about evidence.

Annex 2 describes the data analysis undertaken.

3. DFID 2013 Evidence Survey: The Results

The results of this survey are presented following the narrative set out in the in the conceptual framework. Key emerging messages are highlighted in each section.

3.1 Evidence Value and Use

The extent to which staff value evidence in their work is likely to act as an important incentive to its use. The survey posed a number of questions that were designed to explore perceptions around the value of evidence and the extent to which staff use evidence now and expect to use evidence in the future.

DFID staff believe that evidence can help to improve performance and impact at all points through the programme and policy cycles.

What do DFID staff think about evidence?

Overwhelmingly DFID staff believe that evidence can help to improve performance and impact at all points through the programme and policy cycles. The majority of respondents reported the overall value of evidence to their work as 'A lot' and this was the most common response across grades, divisions, and cadres. With the exception of B1(D)'s, more than 50% of staff chose this response. Only 16 of 446 responses (4%) reported that the overall value of evidence was 'None' or 'A little'.

Evidence – in particular, synthesised evidence – is valuable for challenging assumptions and arguing our case.

Table 3: Rate the overall value of evidence to your work

	SCS	A1	A2	A2(L)	B1(D)	Other	**TOTAL**	**TOTAL (%)**
1 None	1	0	0	3	0	0	**4**	**0.9**
2 A little	0	0	4	5	2	1	**12**	**2.7**
3 Some	3	9	14	6	7	1	**40**	**9.0**
4 Quite a lot	4	34	48	16	4	2	**108**	**24.2**
5 A lot	17	76	139	34	12	1	**279**	**62.6**
Don't know	0	0	0	0	2	1	**3**	**0.7**
Total	***25***	***119***	***205***	***64***	***27***	***6***	***446***	**100**

Generalists reported a significantly lower overall value of evidence to their work compared to those in cadres[4] while respondents from the Research and Evidence Division (RED) reported a significantly higher value of evidence than those from other divisions. Box 1 shows a range of responses that illustrate the values placed on evidence by DFID staff gathered in the focus groups and survey.

[4] An OLS regression was run of the overall value of evidence against dummy variables for cadre, location, division, and grade. The coefficients on all cadre variables were positive, 8 of them significantly so (economic, education, evaluation, health, infrastructure, livelihoods, social development, and statistics). This implies that generalists, the base cadre for the regression, reported a significantly lower overall value of evidence to their work than the majority of advisory cadres.

Box 1: What DFID staff think about evidence

Evidence helps us to challenge the big statements *(FG WH1)*	*Evidence is beneficial in arguing our case with Government* *(FG DRC,I,K)*	*Evidence is central to identifying what works and what doesn't work* *(FG DRC)*	*We think evidence is valuable but can't say for sure that it makes a difference!* *(FG WH1 WH2)*
'There is a reluctance to make judgements about evidence. People are more concerned with making it appear like they are using evidence well but lack confidence in assessing it well.' *(Qualitative survey response)*		*In my opinion, outside the professional cadres, there is less a question of lack of access to evidence and more a lack of interest in actually looking for and using evidence. This can't be entirely explained by a lack of time to look for evidence.* *(Qualitative survey response)*	

When pressed on the basis for this firm belief, focus group respondents were clear that whilst in theory the use of evidence should improve impact, in practice they couldn't say for sure that it did (FG WH1-2). Indeed, 26% of survey respondents reported that they didn't know what impact using more evidence had had on development results and a further 23% thought it had no impact at all. Participants identified a need for DFID to be more rigorous in its attempts to demonstrate the value of evidence use (*table 4*).

We need more proof of how using evidence leads to better development results

Table 4: Please rate what impact (the change in time spent on evidence) has had on delivering development results

	SCS	A1	A2	A2(L)	B1(D)	Other	**TOTAL**	**(%)**
1 Highly negative impact	0	2	0	0	0	0	**2**	**0.4**
2 Negative impact	0	6	10	2	0	0	**18**	**4.0**
3 No impact either way	5	27	45	21	4	1	**103**	**23.1**
4 Positive impact	14	54	85	17	6	3	**179**	**40.1**
5 Highly positive impact	1	9	8	5	0	0	**23**	**5.2**
Don't know	2	20	56	17	17	2	**114**	**25.6**
Total	***25***	***119***	***205***	***64***	***27***	***6***	***446***	**100**

So how much do we use evidence?

Most staff across DFID report an increase in the use of evidence over the past three years

Analysis of the survey data [5] consistently finds a relationship between respondents' overall rating of the value of evidence and the extent to which they rely on evidence now and expect to rely on it in the future. 72.5% of the respondents currently rely on evidence in their work either "A lot" or "Quite a lot". Staff from across DFID also reported an increase in focus on evidence over the past three years, with respondents from RED and the SCS reporting the largest increases (*table 5)*.

Table 5: To what extent has the amount of time and resources invested by staff in your area in using evidence changed over the past 3 years?

	SCS	A1	A2	A2(L)	B1(D)	Other	**TOTAL**	**(%)**
1 Much less than before	0	0	0	0	0	0	**0**	**0.0**
2 Less than before	1	2	1	0	1	0	**5**	**1.1**
3 The same as before	0	14	33	16	3	3	**69**	**15.5**
4 More than before	11	64	93	24	5	2	**199**	**44.6**
5 Much more than before	9	28	27	3	0	1	**68**	**15.2**
Total	***21***	***119***	***205***	***64***	***27***	***6***	***446***	**100**

Both the DFID evidence survey 2013 and World Bank 2011 survey included questions on the extent to which staff rely on evidence currently and expect to rely on evidence in the future. The results of the comparison are described in box 2 and show that DFID staff report that they tend to rely on evidence in their work to a greater extent than staff in the World Bank.

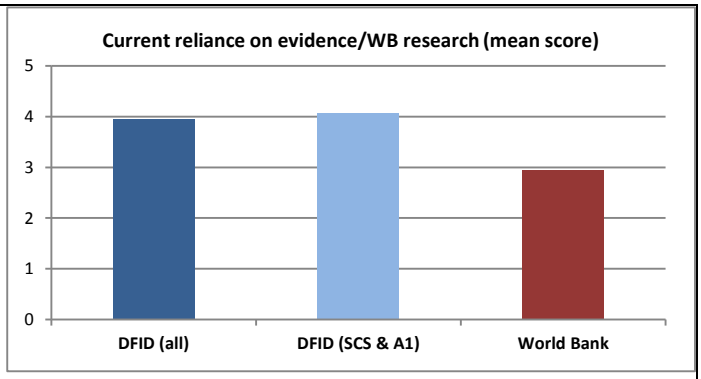
Has this all been worth it?

Analysis [6] of whether staff feel that this extra investment in time has influenced development impact shows that greater perceived increases in the amount of time and resources invested by staff in the use of evidence tend to be associated with reporting a more positive impact on development results.

5 An OLS regression was run of overall value of evidence against current reliance on evidence and controls for location, division, grade, and cadre. The coefficient on reliance on evidence was significantly different from zero (p-value< 0.001), positive, and of a similar order of magnitude to the correlations found in the non-parametric tests (Spearman's Rho and Kendall's Tau-)

6 OLS regression of the change in time and effort devoted to the use of evidence on the perceived impact of this on results and controls for cadre, division, location, and grade also found a significant positive relationship. The coefficient on perceived impact on results is of a similar order of magnitude to the correlation found in the non-parametric tests.

Box 2: DFID and the World Bank[7]



Question 8 in the DFID survey asked respondents to 'rate the extent to which you currently rely on evidence for your work', with responses ranging from 1 (none) through to 5 (a lot), and an option for 'don't know'. Responses to the World Bank survey were collected on a 10-point scale (with 10 being 'very much'). In order to make the means directly comparable, a transformation is applied to the World Bank results.[8] However, it should be noted that the population for the DFID survey was slightly wider, as the World Bank survey population included only senior staff (accounting for approximately one quarter of Bank staff), and that the World Bank question related to World Bank evidence only, whilst the DFID survey related to evidence in general. Therefore, caution should be exercised in drawing comparisons between the results. We present results for all DFID staff and DFID SCS & A1 (who comprised 33% of respondents).

	Current and expected reliance on evidence					
	Current reliance on evidence/WB research			Expected increase in demand (score for expected future demand-score for current)		
	Mean	St. Error	*N*	Mean	St. Error	*n*
DFID (all)	3.95	0.05	446	0.46	0.03	443
DFID (SCS & A1)	4.07	0.07	144	0.40	0.06	144
World Bank	2.95	0.06		0.55	0.03	

With regards to current reliance on evidence, the (adjusted) mean for DFID is significantly[9] higher than that of the WB, implying that DFID staff tend to rely on evidence for their work to a greater degree than their WB counterparts rely on WB evidence. This difference is even more stark when we restrict the sample to only senior DFID staff.

The expected change in reliance on evidence is calculated as the difference between expected future reliance on evidence and current reliance. Both surveys indicated that respondents expect to rely on evidence more in the future than they do currently, but the World Bank survey showed a larger expected change than that for DFID. However, this difference was not statistically significant for either DFID sample.[10]

---

[7] See Ravillion, Martin (2011). "Knowledgeable Bankers? The demand for research in World Bank operations", World Bank Policy Research Working Paper 5892.

[8] The transformation performed is *[(y-1)/2-0.25]+1=Y*, where *y* is a value on the World Bank 10-point scale and Y is the transformed number. Effectively, this means that the two highest values on the WB survey map to the highest value in the DFID survey, the two next highest values in the World Bank survey map to the second highest value in the DFID survey and so on. The standard errors on the World Bank survey are divided by 2, whilst those on the DFID survey are left unchanged.

[9] T-tests were conducted for differences between the means. This involved making assumptions about the number of responses to this particular question in the World Bank survey. The overall response number is reported as 555, so tests are assuming that all respondents answered this question, and as a robustness check, assuming only half of overall respondents answered this particular question (278). The inference from both tests was consistent.

[10] Again, t-tests were run under different assumptions for sample size in the World Bank survey. Both tests assumed responses of 555 and 278 resulting in a statistically insignificant difference from the DFID survey (for both DFID samples).

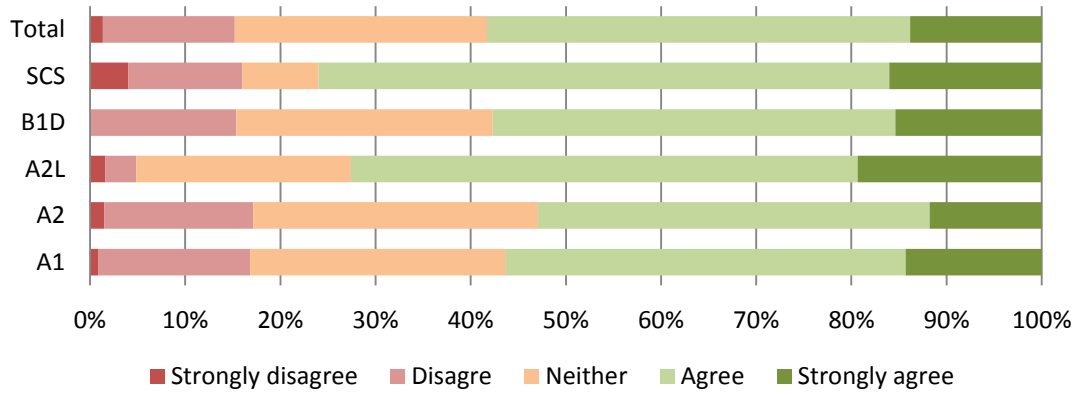
3.2 Management Support for the Use of Evidence

This section explores the external factors that incentivise the use of evidence. In an ideal scenario, staff would both value evidence and be operating in an environment that encourages its use. However, reality may not match this ideal. Consequently, the survey included a range of questions that focused on the perceptions of DFID's values on evidence and lesson learning as well as managerial support in both planning and priority setting processes.

"..DFID is not solely driven by evidence because we are not working in a vacuum.." (FG WH2)

58% of DFID staff agreed or strongly agreed that their managers value and scrutinise the use of evidence. Considered by grade, A2(L) and SCS had the largest proportion of respondents who answered positively.

Graph 1: Proportion of respondents who feel their managers value evidence and scrutinise its use



Do senior managers place less emphasis on evidence?

One might expect that as staff move up through the grades (and away from roles that have a majorly technical focus) the range of factors that influence their priorities and decisions increase and thus the relative influence of evidence decreases. This view is offered some support by the qualitative survey responses and in the focus groups. Box 3 gives an illustration of the range of comments that we found. These relate both to the competing drivers in decision making process and how these affect the reality of evidence use.

However, the survey responses presented a more complicated picture. The survey results show a difference between staffs own values for evidence and their assessment of the value placed on it by their managers. This difference is large, as illustrated by the following stylised example, which is based on the assumption that B1D responses to the question "Do managers value evidence and scrutinise its use?" (MVE) refer directly to A2s, and A2s to A1s and A1s to SCS. We compared each grades assessment of their own value of evidence with the perceptions of the grades below them. This results in the following table, which firstly shows the actual scores for each grade to the two questions and then it shows the differences described above.

Table 6: The difference in value on evidence between staff and their managers

Grade	SCS	A1	A2	A2(L)	B1(D)	Total
Own value of evidence (a lot)	68%	64%	68%	53%	44%	63%
MVE (strongly agree)	16%	14%	12%	19%	15%	14%
Grade (manager – staff)		**SCS - A1**	**A1 - A2**		**A2 - B1(D)**	
Difference (percentage points)		54	52		52	

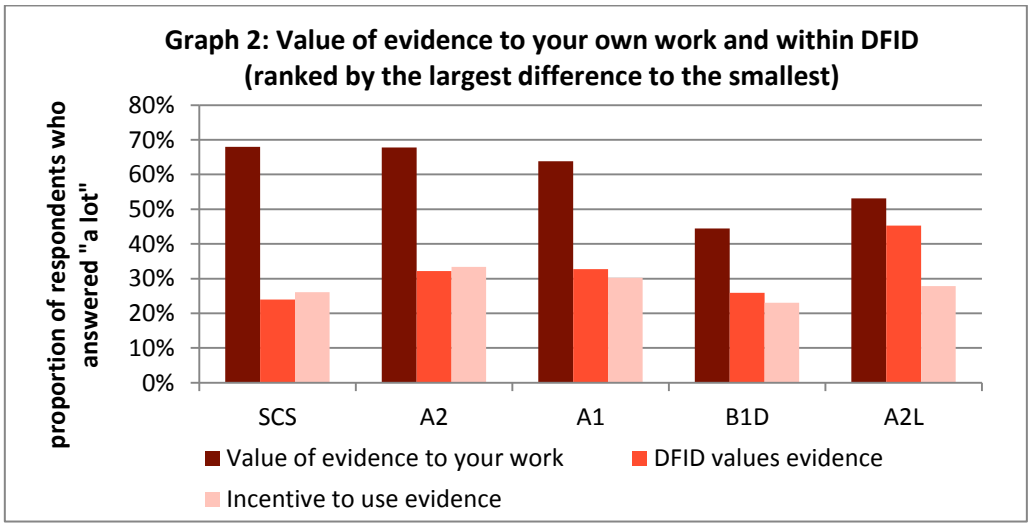
For the SCS respondents and their A1 staff there was a 54 percentage points difference in the results. This level of difference is consistent across the stylised grade hierarchy.

It should be noted that the "managers value evidence" scores may be particularly low because the question also includes "… and scrutinise my evidence" so the difference may be capturing a feeling of a lack of scrutiny and not just the sentiment that they place little value on evidence.

Does DFID as an organisation value evidence?

In a similarly mixed picture, the survey also found a large disparity in staffs' perceptions of the value of evidence to their own work and the value of evidence to DFID as an organisation. This difference is largest with SCS at 44 percentage points and tends to narrow as you move down the grades, with the smallest difference with B1(D)s and A2(L) at 19 and 8 percentage points, respectively.

Graph 2: Value of evidence to your own work and within DFID (ranked by the largest difference to the smallest)



This disparity occurs across the grades and so could be seen to suggest that staff may be misinformed about the importance that other staff place on evidence, or we may be observing some behavioural biases - something akin to an "illusory

superiority bias".[11] However, this depends crucially on whether staff consider "DFID as an organisation" to mean "other staff within DFID" or something closer to "the incentives that exist within DFID", which are related but not identical. Given how similar the scores are for "incentive to use evidence" and "DFID values evidence", staff may have viewed it as more of the latter.

Box 3: Quotes from survey question 24 "Is there anything else you would like to say about evidence use in DFID" and themes from the focus groups

"The culture and processes have improved a lot - but we need more consistent incentives to keep it front and centre of better development. Often people talk about political trade-offs but it's usually lack of integrity or sloppiness that impedes use of good evidence married with political savvy." (qualitative survey response)

"The message on use of evidence needs to be reinforced at all levels, but at Deputy Director level they are not doing this and consequently teams feel it's OK to prioritise anything over improving evidence." (qualitative survey response)

"Where the evidence agenda clashes with other agendas (such as the pressure to spend more money) then evidence often comes out in second place." (qualitative survey response)

"I think we have to be realistic about the relative weight between evidence and politically driven decisions. Can we be clearer about difference situations in which one or the other of these holds the greater strength?" (qualitative survey response)

Value placed on evidence needs to extend beyond the business case (FG I, K).

Do staff have enough time?

If evidence is a priority it needs to influence the very basic blocks of individual and collective work plans.

The answer to this question may legitimately be "when do we ever have enough time!" but a very strong response across the survey showed 44% of respondents either "completely disagreed" or "disagreed" with the statement that "I have enough time to consider evidence". The qualitative survey responses highlighted that time pressures within DFID make it difficult to access, absorb and apply evidence (25 comments). Specific concerns included the pressure to spend quickly, as well as business processes such as annual reviews and business case development and policy briefings where demands for a rapid turnaround restricted the ability of staff to access evidence. As such, staff feel that they have insufficient time to digest and even less time to disseminate relevant new evidence.

[11] Examples of this bias include Svensons (1981) finding that 93% of a US sample and 69% of a Swedish sample put themselves in the top 50% (above the median) in terms of driving skill.

In addition, one of the themes that came across in both the focus groups and the survey's qualitative responses (35 comments) was that the use of evidence is often confined to the business case process. Focus group participants acknowledged that ideally evidence should be used at different stages of the policy and planning processes.

These all highlight that where a priority on evidence is stated, the way that this translates into business planning processes and individual work planning will have a significant influence on whether that priority is realised.

3.3 Evidence skills and knowledge

Staff need the right skills and knowledge to be able to identify and find the evidence they need; assess how reliable its findings are; understand how it fits into a wider body of evidence; and apply it effectively to their work.

Do DFID staff have the necessary skills to use evidence effectively?

Staff confidence in their skills to find and use evidence varies significantly between grades and cadres.

The survey asked staff to agree whether they had sufficient skills to find and use evidence well. Staff were then asked to rate their specific skill levels on the four main types of evidence (research, evaluation, qualitative, quantitative). The survey also asked staff to list what training they would find helpful to further build these skills.

Overall, two thirds of staff agreed or strongly agreed that that have the necessary skills to find and appraise evidence[12], with only 14% of staff disagreeing. Self-assessed skills ratings were broadly similar across all four main types of evidence.

Table 7: Please rate your own skills in each of the following areas:

	Finding and appraising research evidence		Finding and appraising evaluation evidence		Using and interpreting quantitative data/statistics		Using and interpreting qualitative data	
1 Complete beginner	19	**4.3%**	21	**4.7%**	22	**4.9%**	14	**3.1%**
2 Beginner	66	**14.8%**	70	**15.7%**	71	**15.9%**	54	**12.1%**
3 Intermediate	168	**37.7%**	199	**44.6%**	175	**39.2%**	174	**39.0%**
4 Advanced	148	**33.2%**	118	**26.5%**	136	**30.5%**	158	**35.4%**
5 Very advanced	38	**8.5%**	29	**6.5%**	38	**8.5%**	40	**9.0%**
N/A	7	**1.6%**	8	**1.8%**	4	**0.9%**	4	**0.9%**

[12] This is corroborated by weak but statistically significant correlation between people who said that they had the right skills to use evidence and their ability to answer technical questions.

However, this reported level of confidence varied significantly between professions, grades and locations. In particular:

- Professional cadres[13] reported much higher confidence levels than generalists, with 68% of advisers agreeing or strongly agreeing that they had sufficient skills, compared to 40% of generalists.
- Staff confidence in their skills to use evidence was broadly consistent across grades B1(D) to A1, but lower for SCS (48%).
- Self-reported skills varied slightly by team, with the Research and Evidence division and West and Southern Africa division having the highest proportion of staff that agreed or strongly agreed that they had sufficient skills.

In addition to staff perceptions of their own skills a series of technical questions were set to give indication of actual knowledge. The answers to these questions are shown in Annex 1 and a breakdown of the answers by grade and cadre is given in Annex 3.

What training do DFID staff want?

Staff want more training on a range of subjects, from basic critical appraisal to more complex comparative analysis.

Qualitative survey responses highlighted a number of key areas where staff wanted more training support to develop their skills. This included: finding and critically appraising evidence (76 comments); statistical evidence, including econometric or quantitative data interpretation (75); evaluation skills (20); qualitative data (19); research skills (12) and applying evidence in policy (11).

Focus groups highlighted further training needs in using evidence to appraise options; assessing the quality of evidence; and applying evidence from other contexts or with a different focus. They also raised the concern that it is unclear what level of skills are needed to be a "critical consumer", which can lead to a belief that developing a high level of competency is out of reach of most staff.

> *"…There is a perception that you have to enter into the world of academia to fully understand research findings. This problem is exacerbated by the fact that research can be both 'faddish' and 'nuanced'…." (FG WH3)*

How should this training be delivered and to whom?

Staff want more training that is practical, tailored to different needs and accessible for country offices.

When further explored through focus groups, staff reported that whilst they value the training support which is available, the current offer is seen as too limiting, particularly for country office staff. Without a focal point for organising training courses, the danger is that training can either end up as UK focused and poorly adapted to country contexts; or organised on an ad-hoc basis and therefore dependent on staff with the right skills and enough time to organise bespoke courses. Box 4 highlights the ideas expressed by staff in both the focus groups and survey qualitative responses around ways that training can be improved.

[13] The economics, health, social development, and statistics cadres had higher self-reported skills

Box 4: How can we better support staff to build their skills

- Provide practical training on how to use evidence for DFID's work, for example using evidence to appraise options in a business case.
- Tailor courses to different levels of need and roles within the organisation. This could include courses aimed at programme managers as well as advisers, and refresher courses as well as full training packages.
- Provide a better understanding of how generalists can and should use evidence – "…there is a perception that evidence is just something for advisers, which is perpetuated by adviser retreats…as opposed to little training available for generalists…" (qualitative survey response).
- Expand the training opportunities for country offices, with a focus on training that is case specific and tailored to the country context.
- Include 'practical training', for example shadowing a research team in the field to gain an understanding of data collection challenges. "…This approach would be less abstract than office training and improve knowledge and experience…." (FG DRC)

3.4 Evidence Availability and Accessibility

Do people know about, use and value the various resources to access evidence?

DFID provides a number of services and tools to enable staff to review and use evidence in their work. Staff were asked to rank their perception of the barriers to using evidence in their work, and then to rate the usefulness of centrally provided DFID evidence services. Staff belonging to a professional cadre were also asked to rate the effectiveness of their Chief/Head of Profession (CoP/HoP) as well as the extent to which their cadre network helped them to navigate evidence.

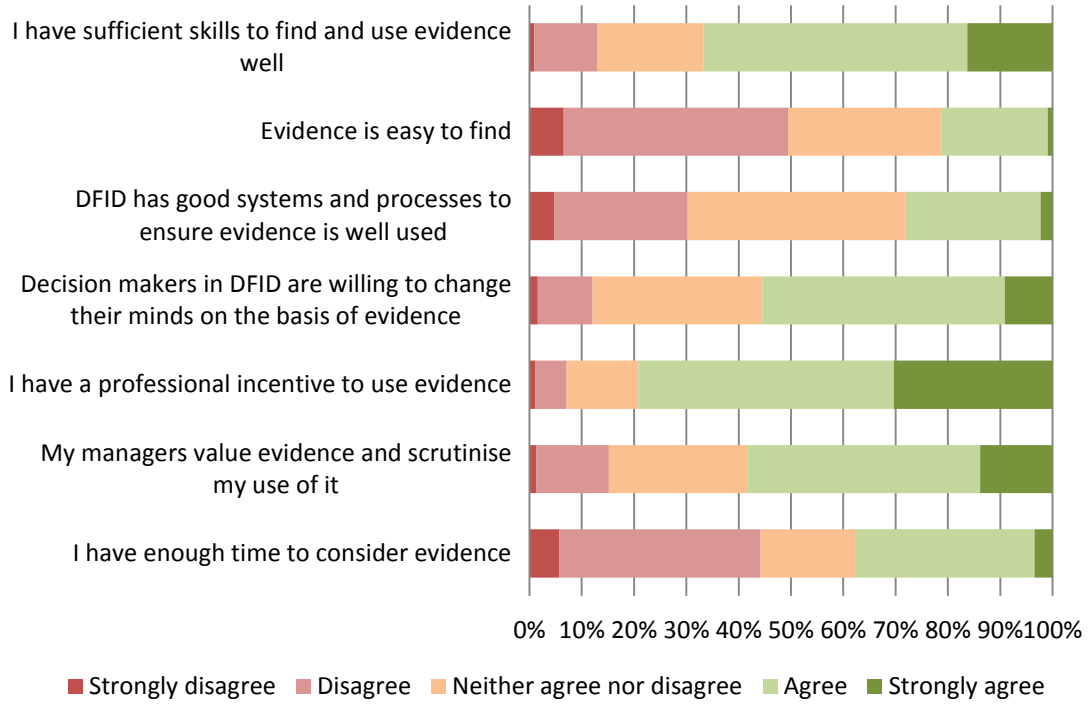
What are the barriers to using evidence?

Staff identified two main barriers to using evidence in DFID; that of easily finding evidence[14]; and that of having enough time to consider evidence[15]. Encouragingly, staff agreed that there was a strong professional incentive to use evidence, and that their skill levels were not a constraint to using evidence.

The biggest barriers to using evidence are easily finding it and having enough time to consider it.

[14] 44 % disagreed or strongly disagreed with the statement that "evidence is easy to find".
[15] 43 % disagreed or strongly disagreed with the statement that "they have enough time to consider evidence".

Graph 3: Staff were asked to what extent they agreed with the following questions:



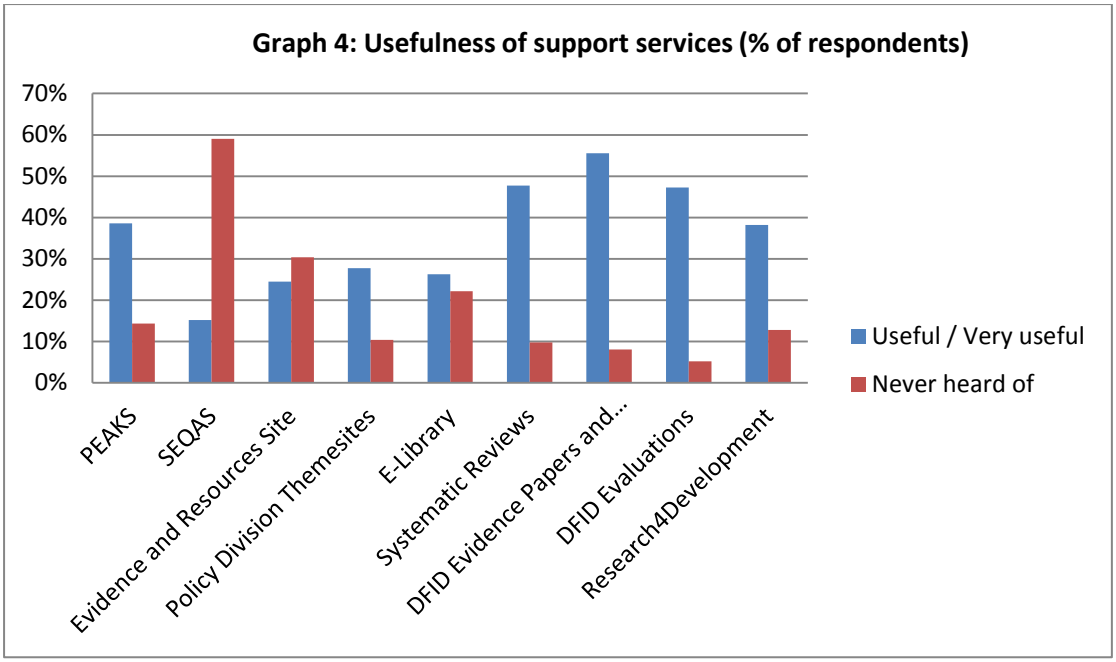
The qualitative survey responses supported these findings, with a large number of respondents suggesting making it easier to access evidence within DFID could help to overcome these constraints (72 comments). In particular, respondents commented that existing evidence resources should be better publicised. Staff also criticised DFID's knowledge management tools, saying that it was difficult to search and find relevant information.

Focus group attendees also gave consistent messages that accessing information in a meaningful way is difficult (FG WH1, K, I), and that any guidance on the use of evidence needs to reflect the amount of time that people have (FG WH1, WH3).

How useful do staff find DFID services to accessing evidence?

DFID services which synthesise evidence are valued but their profile needs to be raised so that staff know they are available

Staff had contrasting views when asked whether DFID has good systems and processes for evidence, with 30% disagreeing or strongly disagreeing whereas 28% agreed or strongly agreed. When asked to rate the tools provided centrally by DFID, staff identified those that synthesised evidence as the most useful. These included DFID Evidence Papers and Literature Reviews, Systematic Reviews and DFID evaluations. External platforms funded by DFID such as the Professional Evidence and Applied Knowledge services (PEAKs) and Research 4 Development (R4D) were also rated as useful.



The messaging from the focus groups is consistent with these findings and attendees highlighted that "…signposting and summarising research [was] helpful, partly because it makes the evidence base much easy to sift and interpret…" (FG WH1, WH2, WH3, K). However there is still room for improvement, with survey respondents requesting better and more consistent communications about the services on offer (45 comments). Respondents also commented that needing to complete a business case for either PEAKs support or to gain access to a journal via the eLibrary was a deterrent to their usage. The focus groups also highlighted that the narrow focus of synthesised evidence products such as systematic reviews could reduce their applicability for programmes, and that their value lessened if not kept up to date (FG WH1).

Internal information systems make it difficult to for staff to access relevant information easily.

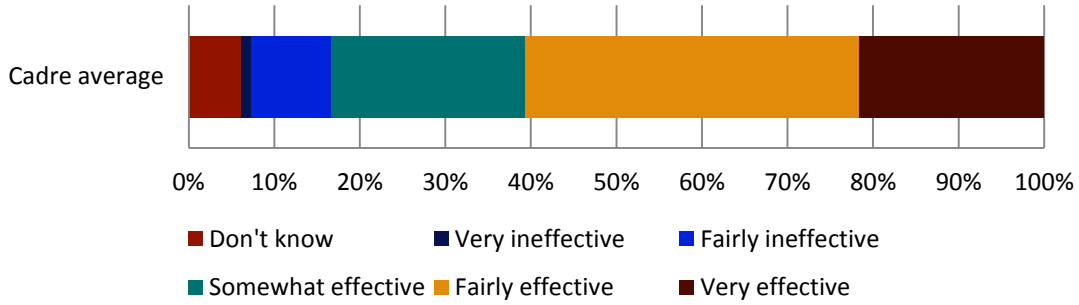
The poor accessibility of these services was highlighted by both the focus groups and the qualitative responses. This is reflected in the survey data where information portals such as the 'Evidence and Resources site' and 'Policy and Division Themesites' rated poorly for usefulness to DFID staff. The focus groups commented that internal information management systems made it difficult to access information (FG WH1, WH2) and the qualitative survey responses suggested a central hub with improved searchability and tagging in the system to relevant projects and programmes (52 comments). However, it should be noted that since the survey was carried out the 'Evidence and Programme Exchange' (EPE) has been launched on inSight, as a 'one-stop-shop' for evidence.

How useful do staff find DFID networks to access evidence?

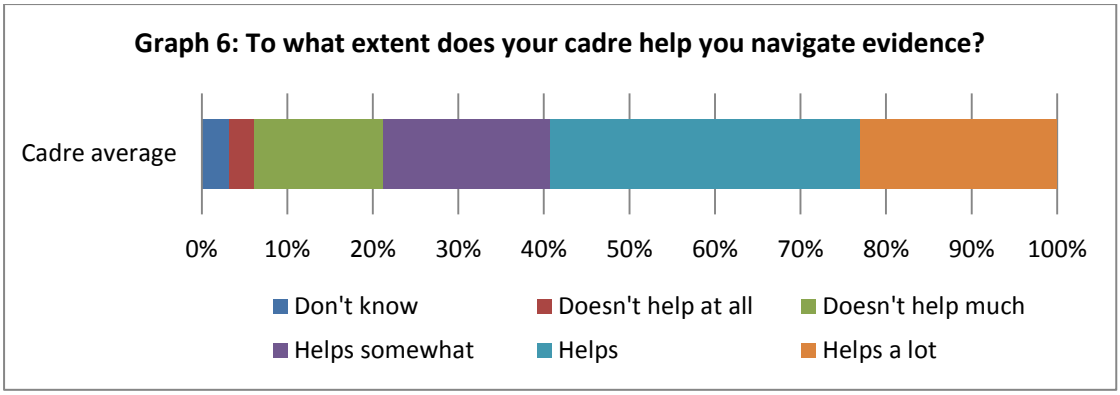
Cadres and informal networks area are a great source of knowledge and experience in DFID. But staff would welcome a more coordinated approach to information sharing.

Staff identified that the personal and professional networks within DFID are important ways in which staff can access evidence. Over 60% of staff belonging to a professional network found their CPO/HoP to be at least fairly effective in marshalling emerging evidence, innovation and new practice or policy. The Economics, Social Development and Livelihoods cadres in particular rated their CPO/HoP as effective. For cadres where there had been frequent changes in personnel, or a recently appointed CPO/HoP, the perceived effectiveness of the HoP decreased. Ways in which HoPs could improve their interaction with their networks was documented in the qualitative comments. These included suggestions for more regular structured communications about the latest evidence or more regular learning events in addition to the annual Continual Professional Development (CPD) conference.

Graph 5: Do you feel that your Chief Professional Officer and Head of Profession are effective thought leaders?



Similarly, respondents reported that their professional networks were helpful in navigating evidence, with over 80% of respondents finding them at least somewhat helpful, and over 60% finding them at least helpful. Advisers belonging to the Statistics, Social Development, Livelihood, Health, Evaluation and Education cadres found their professional networks were more helpful (in navigating evidence) than the average respondent. The qualitative survey responses supported this finding, as many respondents commented positively about their cadre (41 comments). The focus groups also identified networks of colleagues and cadres as very important for accessing evidence quickly (FG WH1, WH2, WH3).



However, advisers also reported that that the informal set-up of the networks could also be a barrier as information sharing is ad-hoc. To overcome this, respondents suggested using a more systematic or coordinated approach to disseminating information that included everyone within a network.

Box 5: Improving access and availability of evidence

- Better communications about what services and support are available: *'…regularly advertising online resources and evidence products helps increase awareness and uptake, particularly within country offices…'* (FG K, I)
- Encouraging staff to use the Evidence and Programme Exchange (EPE) on the DFID inSight pages, including working with the HoPs to publicise its usefulness to cadres.
- Working with HoPs to emulate best-rated cadre networks with regular communications and learning events about the latest evidence.
- Encourage more formal cross-cadre information sharing to widen networks, including shared online repositories for sharing information and evidence resources.
- Championing best practice for sharing information across DFID such as the Evaluation Digest (www.scoop.it/t/evaluation-digest).
- Roll-out of critical appraisal training for the use of evidence by the Evidence into Action team.

3.5 Research Design

Do research, evaluation and other teams involve other business units sufficiently in the design of research and evaluations?

RED commissions thematic research for the global public good via its research teams, and encourages the uptake of research programmes through its Evidence into Action team. The Evaluation Department (also part of RED) oversees evaluation policy, supporting operational staff in commissioning evaluations and promoting the

use of evaluation findings. Staff were asked whether they felt sufficiently engaged with RED in the design of research and evaluation.

How engaged are staff with centrally commissioned research?

Overall staff are "somewhat satisfied" with centrally commissioned research but more can be done to engage country offices and link research to policy demand.

Overall staff reported feeling at least "somewhat satisfied" with centrally commissioned research (57%). But a significant proportion of staff reported that they were dissatisfied or very dissatisfied (18%) and more than 1 in 5 staff (23%) stated that centrally commissioned research was not applicable to their work. When the response was considered by grade, over 30% of SCS reported that they were dissatisfied with central research, and A1 staff were more likely to be dissatisfied than A2 staff. There was no observed significant difference in satisfaction across cadres. Country office-based staff also reported slightly lower satisfaction levels with 57% at least "somewhat satisfied" and 20% at least dissatisfied.

Table 8: Please rate your satisfaction with your level of involvement with centrally commissioned research

	SCS	A1	A2	A2(L)	B1(D)	Other	**TOTAL**	**TOTAL (%)**
1 Very dissatisfied	0	4	1	1	1	0	**7**	**1.6%**
2 Dissatisfied	7	30	27	6	3	0	**73**	**16.4%**
3 Somewhat satisfied	7	43	75	19	5	2	**151**	**33.9%**
4 Satisfied	5	17	55	6	3	0	**86**	**19.3%**
5 Very satisfied	0	11	4	1	2	0	**18**	**4.0%**
N/A	3	14	41	28	13	4	**103**	**23.1%**
Total	***25***	***119***	***205***	***64***	***27***	***6***	***446***	

The focus group discussions raised the concern that research findings are not relevant enough to the reality on the ground. Attendees commented that they wanted evidence that was easier to apply to specific programmes and country contexts. In particular, given the length of time required to commission research within a country office, attendees suggested that more regionally relevant research be commissioned. This message was reflected in the qualitative survey responses, where a number of the responses (47 comments) suggested that research be linked with demand in either a country or policy context.

The qualitative survey responses also revealed a contrast between respondents who felt that they had been involved well by central research, and those that felt that they had been excluded from potentially useful research. Further suggestions included improving the communications or dissemination of research projects (36 comments) and encouraging RED to engage more widely on its research portfolio and commissioning process (34 comments).

Box 6: Improving research design

- More consistent communications from RED about the DFID research portfolio. The C/HoPs could be an important link here to disseminate information to cadres and advisers.
- Wider engagement in the research commissioning process to address overlaps between 'global good' need and country, region or policy demand.
- Better links between research and evaluation.

Is evaluation important to DFID?

Staff recognise that evaluation is important but want more support and for findings to be shared widely.

It is encouraging that staff recognised evaluation's importance to DFID, with 80% of staff reporting that evaluation was at least "quite important". Similar levels of staff reported that the support available to carry out evaluations was somewhat useful (72%), with 42% reporting that the support was at least "quite useful". The qualitative survey responses gave suggestions for improving the support offer for evaluation (*box 7*).

Box 7: How can we better support staff in evaluation

- Clarity on what teams are expected to do on their own and when they should seek expert advice from the Evaluation Department.
- Clearer institutional guidance on evaluation approaches and procurement frameworks for commissioning evaluations including guidance on likely costs.
- More access to thematic evaluation advisers from country or policy teams to 'help guide lead advisers or programme managers' and practical support on how to run an evaluation.
- Better communications or dissemination to enable the findings of evaluations to be shared, and one respondent commented that "lessons could be taken from the dissemination of information and tools via the Africa Division Evaluation".
- Training, especially for programme management staff.
- More support on impact evaluations.
- Include evaluation advisers in programme design.
- A stronger link between the monitoring function of programme managers and evaluation.

How well do we learn the lessons?

Learning and applying lessons was another area tested by the evidence survey. Just over half of DFID staff thought that lesson learning in their area was 'average'. The remaining responses were split evenly between staff that thought DFID learnt and

applied lessons ‘well or very well’ (26%) and those that thought DFID did this ‘badly or very badly’ (20%). The focus groups reinforced this latter message suggesting that DFID doesn’t make the most of information and lessons generated by DFID programmes.

Suggestions for improving lesson learning across DFID were documented in the qualitative survey responses. These included a systematic cross-DFID approach to recording lessons (24 comments); wider dissemination and communication about lessons learned from programme reviews and evaluations (31 comments); and training on the use of evidence throughout the policy and programme cycles (20 comments).

4. What Happens Next?

The evidence survey was intended to help DFID to become even better at using evidence to underpin our policy and programming decisions. Through identifying both strengths and weaknesses in our approach to evidence, the survey highlighted a number of key areas where we could do more, including better training, improved accessibility, more consultation and a continued shift in culture.

The report's findings will be used by the Evidence into Action team, Heads of Profession and other Research and Evidence Division colleagues to take action in some of these key areas, including:

- Creating targeted training and support to staff, to respond to gaps in skills and knowledge;
- Developing and updating guidance and resources to help staff use evidence well;
- Informing the continued improvement of the Evidence and Programme Exchange and the review and redesign of the Research 4 Development database; and
- Providing a baseline, for the first time, to judge the progress of the drive to improve our use of evidence.

Depending on demand for the information, other survey results products will include:

- Management briefing notes;
- Data summaries for cadres and divisions;
- Bespoke summaries for specific challenges; and
- Presentations and seminars.

However, these changes cannot be made by RED alone, but will require the continued support and commitments from all DFID staff to placing evidence at the heart of our work.

ANNEX 1: Evidence Survey Questions

Introductory questions

Questions 1 to 7 gathered information on survey participants e.g. profession, grade, location, sex etc.

Main survey

This section of the survey was preceded by the following statement:

"For all questions, unless otherwise specified, 'evidence' should be thought of in terms of evidence from research, evaluation and statistical data sets"**.**

8. Please answer the following questions *(Rated from none (1) to a lot (4))*

- Rate the overall value of evidence to your work
- Rate the extent to which you currently rely on evidence for your work.
- Rate the extent to which DFID as an organisation values evidence.
- Rate the extent to which you expect to rely on evidence in the future for your work.

9. This question looks at the constraints you might face in using evidence. To what extent do you agree with the following questions? *(Rated from strongly disagree (1) to strongly agree (5))*

- I have enough time to consider evidence
- My managers value evidence and scrutinise my use of it
- I have a professional incentive to use evidence
- Decision makers in DFID are willing to change their minds on the basis of evidence
- DFID has good systems and processes to ensure evidence is well used
- Evidence is easy to find
- I have sufficient skills to find and use evidence well
- [Free text] What are some of the key things DFID can do to overcome these constraints?

10. Please rate your own skills in each of the following areas *(rated from complete beginner (1) to very advanced (5))*

- Finding and appraising research evidence
- Finding and appraising evaluation evidence
- Using and interpreting quantitative data/statistics
- Using and interpreting qualitative data
- [Free text] What training on using evidence would you prioritise for yourself?

11. Please rate the extent to which you think that DFID staff in your area have the capability to use evidence... *(rated from very low capability (1) to very high capability(5))*

- In the design of programmes and policies?
- In the implementation of programmes and policies?

12. To what extent has the amount of time and resources invested by staff in your area in using evidence changed over the past 3 years? *(rated from much less than before (1) and much more than before (5))*

13. Please rate what impact you feel this has had on delivering development results? *(rated from highly negative impact (1) to positive impact (5))*

14. How well do you think DFID learns and applies lessons from its work in your area? *(rated from very badly (1) to very well (5))*

- [Free text] What suggestions do you have for improving?

15. Thinking specifically about evaluation evidence... *(rated from not at all(1) to very (5))*

- How important do you feel evaluation is to DFID's work?
- Is the evaluation support and advice on offer sufficient for your needs?
- [Free text] Why and what could be done to improve?

16. Please rate your satisfaction with your level of involvement with centrally commissioned research *(rated from very dissatisfied (1) to very satisfied (5))*

17. Please rate the usefulness of the following services provided centrally by DFID to help staff find and use evidence to your work *(rated from very useless (1) to very useful (5))*

- PEAKS evidence products and help desk
- SEQAS
- Evidence and resources site
- Policy division themesites
- E-library
- Systematic reviews
- DFID evidence papers and literature reviews
- Evaluation department
- R4D
- [Free text] What can be done to make these more useful to your work?

18. Do you feel that your Chief Professional Officer and Head of Profession are effective thought leaders? This means they effectively marshal emerging evidence, innovation and new practice/policy into the cadre? (In the case of economics and private sector cadre, please use Chief Economist as your Chief Professional Officer). *(rated from not effective (1) to highly effective (5)).*

- [Free text] Please explain why and what they could do differently

19. To what extent does your cadre network help you to navigate evidence? *(rated from not at all (1) to a lot (2))*

- [Free text] Please explain why and what they could do differently

Technical questions

The following questions asked respondents to identify key research, evaluation and statistics concepts.

20. Please choose the most appropriate definition for each of the following research terms:

	Research which examines random effects of intervention in a controlled setting	A method for systematically analysing qualitative data	Research which compares effects of intervention in one group to effects in a control group	A research study which analyses the content of statistical datasets	A research study which summarises results from a number of studies	% correct answers
A. Randomised control trial			X			84 %
B. Systematic Review					X	80 %
C. Content analysis		X				48 %

21. Please select the most appropriate definition for the following evaluation terms:

	Process Evaluation	Randomized Control Trial	Posthoc	Impact evaluation	Don't know	% correct answers
A. Any evaluation that systematically, rigorously and empirically investigates the impacts produced by an intervention, using appropriate designs and methods.				X		93 %
B. An approach that looks at how a programme or intervention rolled out in practice.	X					90 %
C. A study that is designed after an intervention has ended			X			90 %

22. Please identify the correct term for each of the definitions below which relate to appraising evidence:

	Reliability	Rigour	Measurement validity	External validity	Don't know	% correct answers
A. Do the 'things' that the researcher is measuring satisfactorily represent the idea they are investigating?	X					15 %

B. Does the researcher measure those 'things' accurately?			X			31 %
C. Do you think that if the research were repeated in a different locality / country, the results would be similar / identical?				X		71 %

23. Country X has a population of approximately 30 million people, around a third of which are school aged children. It is estimated that 90% of children attend school.

- Please select the most appropriate answer for each of the following questions:

	30m	15m	9m	5m	100,000	2,000	Don't know	% correct answers
A. How many people would you expect to be covered by a national sample survey of people's opinions on the state of education?						X		39 %
B. How many people would you expect to be covered by a population and housing census?	X							77 %
C. How many people would you expect to be covered by the administrative data on education?			X					75 %

24. Is there anything else you would like to say about the use of evidence in DFID?

- [Free text]

ANNEX 2: Evidence Survey Approach to Data Analysis

Survey Quantitative data – The majority of the questions in the survey were 'Likert-style items' asking the extent to which respondent's agreed or disagreed with a statement, for example their satisfaction with a particular service. As such, the majority of the data is ordinal; we can make inferences based on the order of responses ('Very satisfied' is more satisfied than just 'Satisfied'), but not scale (we cannot say that 'Very satisfied' is twice as satisfied as 'Satisfied'). The ordinal survey data was analysed to give the percentage of each type of response per question, and then analysed by different variables such as cadre, grade and location.

Non-parametric analysis was then used to test the correlation between different responses, for example 'are respondents that are amongst the most satisfied with question X also amongst the most satisfied with question Y?'. Spearman's Rank Correlation Coefficient and Kendall's Tau-b Rank Correlation Coefficient methodologies were used. Both test how similar the ordering of responses are when ranked by 2 questions and range from -1 (perfect negative correlation) to +1 (perfect positive correlation), with zero indicating no correlation.

Ordinary Least Squares (OLS) analysis was also conducted, with controls included for the respondent's division, cadre (including generalists), location, and grade to account for variation in the data which could disrupt the correlation tests .[16]. For the largest single groups in the sample (generalist, A2 grade and London-based), dummy variables were not included. Policy Division was also analysed without dummy variables because the division was considered a good central benchmark for the use of evidence. Therefore, significant coefficients indicate significant differences from these states in each dimension tested. Where the Breusch-Pagan test indicated significant heteroskedasticity, robust standard errors were used for inference. Where OLS analysis suggests substantially different inference from the non-parametric tests, this is clearly indicated.

Survey Qualitative Data – The qualitative responses to each question where respondents could respond in free text were consolidated to identify trends. Where the analysis has been included in the report, the strength of the trend is indicated by the number of comments in brackets. There are opportunities for further analysis of the qualitative data set but we do not intend to undertake these at this time.

Focus Group Response - The responses from the focus groups are reported in this report with a notation that indicates which groups raised the issues. For example, the message that "evidence helps to improve policy and performance" was mentioned in 6 groups and so is followed by the notation (DRC, I, K, WH1-3).

[16] It would have been possible to perform non-parametric regression analysis, but in light of the fact that non-parametric tests were already run to establish relationships under stricter interpretations, OLS was used to make interpretation of results more straightforward.

ANNEX 3: Results to the Technical Questions by Grade

	Answer	SCS	A1	A2	A2L	B1D	Other	Total
Question 20: Choose the appropriate definition for each of the following research terms								
A. Randomised Control Trial	20A. Correct	95%	87%	86%	65%	89%	20%	84%
	20.A Incorrect	5%	13%	14%	35%	11%	80%	16%
B. Systematic Review	20.B Correct	86%	89%	84%	59%	59%	40%	80%
	20.B Incorrect	14%	11%	16%	41%	41%	60%	20%
C. Content Analysis	20.C. Correct	48%	52%	48%	46%	37%	0%	48%
	20.C Incorrect	52%	48%	52%	54%	63%	100%	52%
Question 21. Please select the most appropriate definition for the following evaluation terms:								
A. Any evaluation that systematically, rigorously and empirically investigates the impacts produced by an intervention, using appropriate designs and methods.	21.A Correct	95%	97%	94%	87%	96%	33%	93%
	21.A Incorrect	5%	3%	6%	13%	4%	67%	7%
B. An approach that looks at how a programme or intervention rolled out in practice.	21.B Correct	86%	92%	94%	76%	96%	33%	90%
	21.B Incorrect	14%	8%	6%	24%	4%	67%	10%
C. A study that is designed after an intervention has ended.	21.C Correct	90%	92%	92%	79%	96%	50%	90%
	21.C Incorrect	10%	8%	8%	21%	4%	50%	10%
Question 22. Please identify the correct term for each of the definitions below which relate to appraising evidence:								
A. Do the 'things' that the researcher is measuring satisfactorily represent the idea they are investigating?	22.A Correct	10%	16%	15%	14%	22%	17%	15%
	22.A Incorrect	90%	84%	85%	86%	78%	83%	85%
B. Does the researcher measure those 'things' accurately?	22.B Correct	19%	29%	39%	21%	19%	17%	31%
	22.B Incorrect	81%	71%	61%	79%	81%	83%	69%
C. Do you think that if the research were repeated in a different locality / country, the results would be similar / identical?	22.C Correct	67%	78%	75%	50%	70%	50%	71%
	22.C Incorrect	33%	22%	25%	50%	30%	50%	29%
Question 23. Country X has a population of approximately 30 million people, around a third of which are school aged children. It is estimated that 90% of children attend school. Please select the most appropriate answer for each of the following questions:								
A. How many people would you expect to be covered by a national sample survey of people's opinions on the state of education?	23.A Correct	48%	43%	44%	13%	33%	0%	39%
	23.A Incorrect	52%	57%	56%	87%	67%	100%	61%
B. How many people would you expect to be covered by a population and housing census?	23.B Correct	76%	82%	81%	62%	70%	50%	77%
	23.B Incorrect	24%	18%	19%	38%	30%	50%	23%
C. How many people would you expect to be covered by the administrative data on education?	23.C Correct	76%	85%	73%	60%	81%	50%	75%
	23.C Incorrect	24%	15%	27%	40%	19%	50%	25%