Executive Summary

Process Evaluation of the International Initiative for Impact Evaluation (2008-11)

27 September 2012

James Morton (Team Leader)

Louise Shaxson (ODI)

Jonathan Greenland

Specialists:

Laetitia Antonowicz (Education)

Simon Armstrong (Private Sector Development)

Felicity Proctor (Agriculture)

Sarah Herbert-Jones (Health)

Nicholas Freeland (Social Protection)

Professor Nicholas Mascie-Taylor (Statistics)

Triple Line Consulting Ltd
3 Princeton Court
55 Felsham Rd
LONDON SW15 1AZ, UK
Tel: +44 (0) 20 8788 4666
3L@tripleline.com
www.tripleline.com

Overseas Development Institute
203 Blackfriars Bridge Road
London
SE1 8NJ, UK
Tel:+44 (0)20 7922 0300
odi@odi.org.uk
www.odi.org.uk

Table of Contents

1 INTRODUCTION 3
2 PROGRESS TO DATE 4
2.1 Progress to date: Impact Evaluation and Systematic Review (2.1 – 2.3) 4
2.2 Progress to date - Policy Advocacy, Communications and Feedback (2.4) 5
2.3 Progress to date: Organisation, Staffing and Management Systems (Ch 2.6 - 2.8) 6
2.4 Progress to date – Funding Status (2.6 - 2.8) 8
3 GOVERNANCE, DIRECTION AND ACCOUNTABILITY 9
3.1 Governing documents 9
3.2 The Membership 9
3.3 The Board of Commissioners 10
3.4 Accountability in practice 11
4 STAKEHOLDER CONSULTATION 11
5 PROCESS EVALUATION 12
5.1 3ie Expenditure by activity and category 12
5.2 The grant management process 13
5.3 Policy and advocacy processes 15
5.4 Advancement and IE services 17
5.5 Planning, budgeting, monitoring and reporting 17
5.6 Assessment 18
5.7 Financial processes 18
5.8 Management processes and style 21
6 RELEVANCE AND EFFECTIVENESS (CH 6) 22
6.1 The theory of change 23
6.2 The logical framework 24
6.3 The potential impact of 3ie grants 25
6.4 Enduring questions and themes 30
7 CONCLUSIONS AND RECOMMENDATIONS (Ch 7) 31
7.1 Relevance and effectiveness 32
7.2 Process issues 33
7.3 Strategic challenges 34
7.4 Outline for a revised strategy 35

1 INTRODUCTION

1. Triple Line Consulting Limited has been contracted by the Bill and Melinda Gates Foundation to carry out a Process Evaluation of the International Initiative for Impact Evaluation (3ie). The Evaluation Team started its assignment in April 2012. As well as visits to 3ie's Delhi headquarters and London offices, the core of the evaluation has been an in-depth review of 3ie's outputs, and the record of the processes behind the production of those outputs. This was supported by e-Surveys of 3ie grantees and a broad community of policymakers, donor officials, and evaluation practitioners; and by interviews with 57 key stakeholders, including Board Commissioners. The Team record their appreciation for generosity and openness of the help they have had from all those involved, the 3ie team in particular.

2. 3ie's strategy document sets out a four year programme: 2010 – 2013. It has five components totalling $51.4 million. With over 75% of the total budget, the focus is on commissioning Impact Evaluations to generate new, policy-relevant evidence (Component 1).

Table 1: Budget Allocation by Component

Budget Allocation	
1. Generation of policy relevant new evidence	$40 million
2. Dissemination of synthesized user-friendly evidence	$2 million
3. Building a culture of evidence-based decision-making	$0.8 million
4. Developing capacity to promote, use & undertake Impact Evaluations	$0.6 million
5. Developing 3ie's institutional capacity	$8 million

3. 3ie's origins go back to 2004, when the Center for Global Development (CGD) launched its *Closing the Evaluation Gap Initiative.* After wide consultation, 3ie was registered as a non-profit US corporation in May 2008. Its Founding Document sets out the following vision:

MISSION "… to contribute to the fulfilment of aspirations for wellbeing by encouraging the production and use of evidence from rigorous impact evaluations"

OBJECTIVES

- To provide public goods that promote the generation and use of good quality evaluations in public policy making;
- To help policy makers and researchers plan quality impact evaluation designs when learning opportunities with new programs arise; and
- To promote good quality impact evaluations that address policy questions of enduring interest.

3ie was to be a grant making body, concentrating on activities which were not undertaken by other organisations and which were "*suited to its character as an independent, non-profit institution governed by a diverse voluntary membership*."

2 PROGRESS TO DATE

2.1 Progress to date: Impact Evaluation and Systematic Review (2.1 – 2.3)

4. Since 2009, 3ie has commissioned 69 Impact Evaluation studies (IE) and 15 Systematic Reviews (SR). 10 more IE grants have been awarded under the Social Protection Thematic Window and a fourth SR window is in process. When the current round of grant awards is complete, commitments will total $36.5 million. To create this portfolio, 3ie has processed eight grant windows, 586 expressions of interest and 664 grant applications: a very substantial task.

2.1.1 Open Window Grants

5. The largest part of 3ie's portfolio has been commissioned through open calls for grant applications (Open Windows): 68 IE grants to date. With an average grant duration of three years, 3ie is due to deliver its first substantial output in 2012: 28 impact evaluations. Of these, six were complete by mid-year. Across all three windows, the Evaluation Team estimate 11 OW grants are likely to be overdue. This proportion is not greater than might be expected for a large, complex programme of studies implemented by a wide range of organisations.

2.1.2 Policy Window Grants

6. Policy Windows to encourage demand-led IEs proposed by governments and NGOs is a key part of the 3ie strategy. The first policy window was launched in 2010. 33 applications were submitted from policy makers and eleven approved. However, it proved difficult to develop workable IEs. Only one full grant and two Project Preparation Grants were approved, none of them yet completed. The second policy window is following a different approach, working more closely with policy-makers and implementation agencies, through Demand Generation Workshops, and awarding Policy Window Preparation Grants to allow qualified research groups to develop the grant proposals. PW2 initiatives are now underway in India, China, Ecuador, S. Africa and the Philippines. Good progress is being made, but the new approach is resource intensive and there is some way to go before it can be certain that 3ie has an effective approach to generating demand-led impact evaluations.

2.1.3 Thematic Window Grants

7. The Evaluation Team has been asked to look at the balance between 3ie's open grant windows and a more targeted thematic approach. This question goes back to the Founding Document, which called for 3ie to identify *"enduring questions about how to improve social and economic development programs."* A substantial effort was made to do this, culminating in a set of workshop recommendations to be *"taken into account by the 3ie Board in determining the priority themes and EQs for 3ie's grant program."* One of the four themes identified – child development – is well represented in the portfolio. Nevertheless, most of 3ie's portfolio to date has been commissioned through open grant windows with no explicit emphasis on particular themes, issues or sectors. The first thematic window, the Social Protection Thematic Window, was not launched until October 2011, with DFID support. This represents an important shift towards a more targeted approach, and agreements for 10 grants are currently being completed.

2.1.4 Systematic Review Grants

8. Systematic Reviews take less time than IEs and five grants had been completed by mid-2012. 3ie staff have completed four more internally. The Evaluation Team estimated that nine of 27 grants were running behind. Although the grantees have been slow, there have also been delays in the

review process. In systematic reviewing 3ie has established an important partnership managing grants processes for DFID, and more recently AusAid.

2.1.5 Summary

9. With what was initially a very small team, 3ie has put together a very substantial portfolio of IE and SR grants. With over 80 major grants awarded, \$35.2 million committed and \$19 million disbursed (as of 30 June 2012) it is clear that substantial progress has been made. If the first large batch of IE grants is delivered on time in 2012, 3ie will have a solid grant management achievement to its credit. There has also been progress in developing new approaches. The first thematic window is a move towards a more structured portfolio, and the second policy window should result in IE grants which reflect real demand from LMIC policy makers. The establishment of robust grant management procedures is itself an important achievement.

2.2 Progress to date - Policy Advocacy, Communications and Feedback (2.4)

10. Advocacy and communication is a central element of 3ie's work and this effort started even before its formal establishment, with the launch of the 3ie website in early 2008. Since then highlights have included two major conferences, and the preparation and implementation of a Learning Strategy for Policy Impact. A capable Policy Advocacy and Communications division (PACO) has been formed and 3ie's expenditure on A&C (i.e. Strategy Components 3 and 4) has risen to 12.2% of the 2012 budget: a significant change from the 2010 – 2013 Strategy, which only allocated 2.8% to these components.

2.2.1 Policy Influence and Impact

11. 3ie aims to achieve policy impact at scale through funding policy relevant studies and making the results accessible to policy makers. Steadily greater emphasis has been given to integrating advocacy and communications into the grant process, culminating in the introduction of Policy Impact Plans (PIP) which grantees are now required to include in their impact evaluation plans. Of 79 IEs, 67 now have PIPs. 3ie also works with grantees to promote the results of their completed studies. Apart from five final reports, nine workshops have been supported and 12 briefing papers or newspaper articles. Two studies have been published in refereed journals.

12. 3ie commissioned ODI to facilitate Policy Influence Clinics to bring 3ie grantees together with their policy counterparts: for training in communications techniques, and to work on the PIPs. A Policy Influence Monitoring project has also been commissioned to monitor and analyse stakeholder engagement in 3ie grants and to develop ways to maximise policy impact. In November 2011 3ie launched the Commitment to Evaluation index (C2E), to establish a peer learning mechanism for 3ie's members to share experiences about getting evidence into use.

2.2.2 Publications

13. 3ie has an active programme of publications spanning Working Papers on conceptual and technical issues (15 to date), Systematic Reviews (6), Evidence Matters, and other policy briefs distilling evidence from systematic reviews and other material (29). Available for download, many of these publications are also offered in hardcopy. 3ie also sponsors the new quarterly Journal of Development Effectiveness, with 14 issues to date, and issues an e-newsletter approximately quarterly to an expanding mailing list.

2.2.3 Events

14. 3ie hosted major conferences in Cairo in April 2009 (co-hosted with AfrEA and NONIE), and in Cuernavaca in June 2010. These aimed to raise awareness of Impact Evaluation, to bridge the gap

between researchers and policymakers, and to promote the sharing of experiences. With over 400 participants, these won a good level of appreciation from a broad range of stakeholders. 3ie also supports seminar series in Delhi, London (with LIDC) and Washington (with IFPRI): 40 to date. These allow researchers and practitioners to talk about methodological issues, and to communicate their findings. They also aim to encourage a wide audience to engage with impact evaluation.

2.2.4 Website

15. The 3ie website has been an important communications tool: it publicises calls for proposals, distributes useful publications, and is the base for key resources such as the expert roster and IE database. In June 2012 the website was redesigned to be more user-friendly to policymakers.

2.2.5 Stakeholder Feedback

16. 3ie has been a reflective organisation, making considerable efforts to get feedback from stakeholders at appropriate points. Between grant applicants, conference participants, and associate members, 3ie has carried out five surveys. Applicant surveys drew out useful feedback and the results are fairly presented in 3ie's reports. They seem to have succeeded in creating a dialogue between applicants and 3ie management, which has drafted action points to respond to the criticisms. The same can be said of the two post-conference surveys, with criticisms clearly reported and addressed. With the exception of Associate Members, survey response rates were reasonable if not high.

2.2.6 Summary

17. Since 2009, 3ie's advocacy and communication has achieved a great deal with a small team, producing a wide range of authoritative publications, supporting an effective website and promoting large and small events. A substantial effort has been made to strengthen engagement between researchers and policymakers so as to enhance the policy relevance and impact of impact evaluations. In PACO, 3ie has the effective advocacy and communications unit it needs to support whatever strategy the executive and the Board decide.

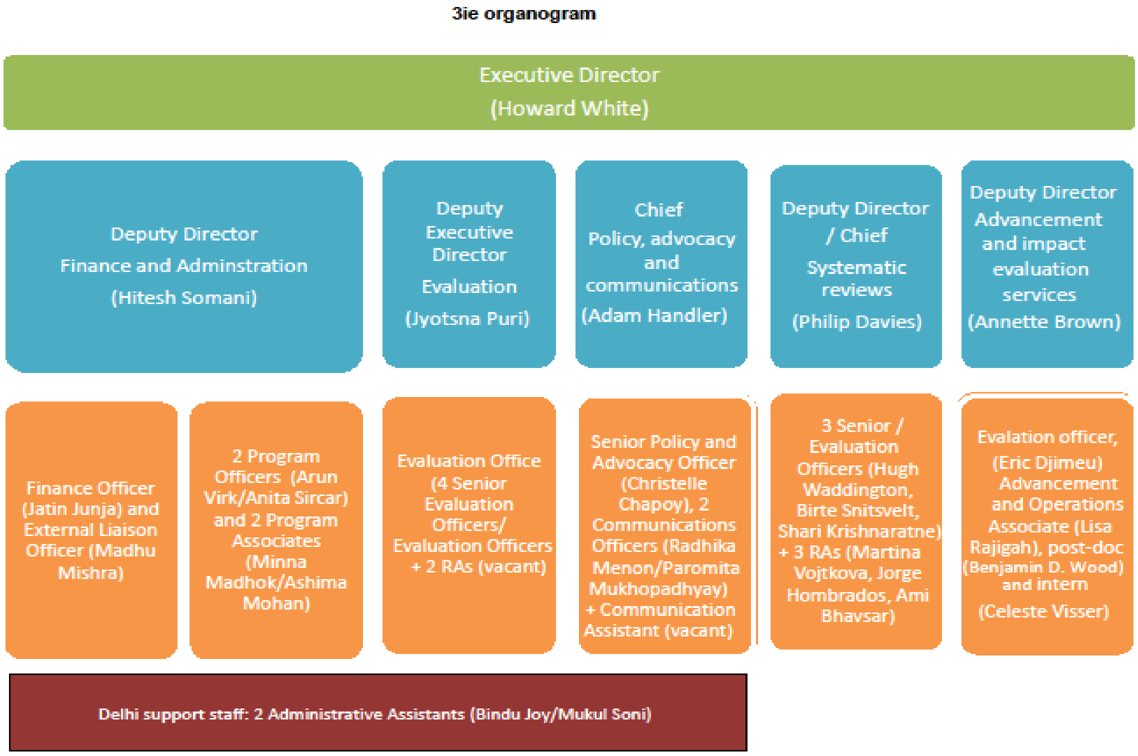
2.3 Progress to date: Organisation, Staffing and Management Systems (Ch. 2.6 - 2.8)

18. 3ie has its five operational divisions: three in Delhi, one in London and one in Washington (see diagram overleaf). The ED is supported by four Deputy Directors (Finance and Administration, Evaluation, Systematic Reviews, and Advancement and Impact Evaluation Services) and a Chief of Policy, Advocacy and Communications (currently being recruited). From a very small start at the beginning of 2009, the 3ie team has grown steadily year on year, adding seven or eight posts each year up until 2011. Until recently, most of the team were in direct contact with the Executive Director or his Deputy. The appointment, between September 2011 and June 2012, of four Deputy Directors is likely to change this balance. 3ie's distributed structure is built around hosting arrangements with GDN, in Delhi, and the London School of Hygiene and Tropical Medicine.

19. Staff in Delhi are recruited on GDN contracts, on GDN terms and conditions. Staff in London are contracted through LSHTM. Only the staff in Washington are contracted directly by 3ie. The hosting agreement with GDN envisages a close partnership between the two organisations, with 3ie forming an International Impact Evaluation Department within GDN. Under the current MoU, GDN provides fully-equipped office space for all 3ie staff in Delhi, telephone services, and one vehicle plus driver, together with the use of GDN's finance, travel and administrative services supported by four dedicated staff. Management of grant contracting and payments to grantees is the much the most important service provided. For this, 3ie pays salaries and other costs as a 'pass through'; a one-off expansion fee for each new staff post; and an annual fee with fixed

element of $600,000 plus 7% of grant disbursements above $4 million.[1] GDN will support a maximum of 22 posts for 3ie.

Figure 1: 3ie Organogram



20. As a membership organisation with strong southern voice, GDN seemed a natural partner for 3ie. Nevertheless, the early years of the relationship were difficult. This has now improved, as 3ie has established its own grant management and financial control functions, giving a clearer division of responsibility between the two organisations. Apart from the provision of offices and backup services, GDN's role is now limited to the issue of contracts and the release of payments to grantees. This is some way from the original expectation that "*3ie activities and programs shall be fully integrated into and form part of the organizational framework of GDN …*"

21. The hosting arrangement with the London School of Hygiene and Tropical Medicine is simpler. Apart from salaries, with a 10% management fee, 3ie only has to reimburse the cost of small items LSHTM buys for it, and an annual office rent; although some of the salaries charged under the agreement seem high. It is understood that the salary/benefit cost of the five Washington staff is $700,000 a year, with office costs of $20,000.

2.3.1 Planning, budgeting and financial control

22. Until late 2010, 3ie had no dedicated finance staff and the Deputy Director Finance was only appointed in September 2011. This was paralleled by the establishment of a stronger Audit and Finance Committee for the Board of Commissioners. This strengthening was completed in April 2012, with the issue of a 3ie Accounting Policies and Procedures Manual. There is no suggestion of financial losses, but prior to that, systems were not strong. At the time of the evaluation team's visit to Delhi, in May 2012, the finance team had just begun to issue monthly statements of 3ie's

[1] 6% above $6 million.

financial position on a cash basis. This is still some way short of a full set of management accounts, by activity or cost centre. Measures are in place to move towards more robust budgeting and financial reporting. However, these will only take full effect from 2013, when divisional Directors will be fully accountable for expenditures against the approved budget. The Accounting Manual does not state it, but it is implicit that it will be the Board of Commissioners' endorsement of a consolidated Workplan and Budget which constitutes the necessary approval.

2.3.2 Human resources

23. 3ie staff are contracted under five different arrangements: international staff contracted through GDN on international terms; Indian staff contracted through GDN on local terms; Delhi staff seconded to 3ie by GDN; UK-based staff contracted through LSHTM; and US-based staff contracted direct by 3ie Incorporated. In principle, Delhi staff are all covered by GDN policies and London staff by LSHTM policies. There have been no major issues to date, but there is some risk that barriers may develop between different categories of staff; although staff retreats and 3ie's internal performance review system should help to overcome this.

24. 3ie is concerned to attract staff of the highest international calibre. Some suggest that few staff in this category are willing to accept a posting to Delhi. This is one reason for maintaining the Systematic Review Office in London, and a Washington Office which is more than just representational. GDN reluctance to create new international staff positions is a complicating factor. To offset this, at least in part, key technical staff members are to be allowed 20% of their time for personal research, although this may need careful management to prevent it weakening the focus on 3ie's own tasks.

25. 3ie has achieved a great deal despite a structure which cannot be ideal for creating a sense of teamwork and common purpose. That it has done so is a credit to the good sense of the staff. Above all, it is the personality, energy and drive of the ED which has been the critical factor in making this structure work. Until now the team has been small enough for most members to have direct contact with him. How to maintain this with a larger management structure will be a key challenge.

2.4 Progress to date – Funding Status (2.6 - 2.8)

26. The Evaluation Report analyses 3ie's funding status on a cashflow basis: actual receipts 2009 – 2012, and committed funding for the period 2013 to 2016. Three points stand out:

- At present only one donor, DFID, has made a commitment to core funding beyond 2013, although it is understood that the Hewlett foundation is about to make a new commitment.
- Three quarters of all core funding has come from just two donors, BMGF and DFID. 90% has come from four: BMGF, DFID, Hewlett and CGD.
- Only one NGO has maintained continuous membership: SCF.

27. Setting this analysis against 3ie's actual expenditures from 2008 onwards, and projections forward to 2016 shows:

- That 3ie will have $6.4 million to carry forward beyond 2016.
- In other words, after meeting current and planned grant commitments and at current operating levels, by the end of 2016 it will have funds for one more year's operation.
- Absent substantial new core funding, 3ie only has surplus core funds for one small grant window of about $2 million.
- Grants make up just less than 60% of 3ie expenditures.

3 GOVERNANCE, DIRECTION AND ACCOUNTABILITY

3.1 Governing documents

28. Three documents govern 3ie: the Certificate of Incorporation, the Founding Document, and the Bylaws. 3ie was registered, in May 2008, as a non-profit corporation in the US state of Delaware, under the standard status for US NGOs and philanthropic foundations. Under the terms of the Certificate of Incorporation, it is the Bylaws which will set out arrangements for 3ie membership, with the rights and benefits of members, and for the regulation of the internal affairs of the corporation, including the Board of Directors.

29. The Founding Document for Establishing the International Initiative for Impact Evaluation (FD) was published in 2008. This sets out 3ie's mission and operating principles, the activities it will undertake, and the governance structure. The structure is headed by a Members' Conference, where full members are development agencies which implement at least $1 million in development programmes annually, and which pay at least a minimum annual fee for membership. The Conference appoints the Board of Commissioners, which is responsible in turn for hiring the Executive Director, approving annual budgets and strategies, and evaluating performance. Commissioners are expected to have a commitment to the use of IE evidence in policy making and to represent major stakeholder groups. Complex arrangements for electing the Board of Commissioners reflect 3ie's vision of a broad membership drawing together donor agencies, their developing country partners and NGOs. As discussed below, this vision has not yet been realised.

30. An Addendum to the FD granted the Executive Director authority to approve contracts and grants according to a graduated scale of purchasing/review procedures. This transitional arrangement has now been replaced by the 2012 Accounting Policies and Procedures Manual. The Evaluation Report notes that the Manual makes no explicit provision for the Board to review and approve budgets or major contracts and grants. In practice, however, the Board does review grant awards as a recommended list, the annual Workplan and Budget, and larger contracts.

31. The 3ie Bylaws have been amended a number of times, with the latest version adopted in April 2012. In most respects they reflect what is set out in the Founding Document. One important change is the decision, in 2011, provide for two Institutional donor seats on the Board, one for Governmental and one for non-governmental donors. With only two meetings a year, Board action is also permitted, subject to the consent of all Commissioners, in writing or electronically; such actions to be filed with the Minutes of the proceedings. There are four standing committees of the Board: Executive, Audit and Finance, Governance and Program, each with its own charter. It can be deduced from the very short list of restrictions that the Board can delegate the majority of its operating powers to the committees.

3.2 The Membership

32. Membership has been a difficult question from the outset. The intention was for southern perspectives to be well represented in 3ie's governance structures and strategy, with donors and developing countries making up 3ie's institutional membership. It was also hoped to give IE practitioners a stake in 3ie's work through non-voting associate membership. Against that vision, 3ie's membership remains disappointingly low. In 2012, it has 22 institutional members: 10 bilateral and one multilateral donor agencies; six developing country governments/agencies; two philanthropic foundations; and three INGOs.

33. 3ie has made concerted efforts to address this issue and the 2012 budget includes increased provision for expanding the membership and funding base. Nevertheless, reviews have suggested the possibility of changing the membership strategy, even to recognise that southern members

are not interested in 3ie's governance and seek a non-membership model. The Evaluation Report suggests that 3ie's emphasis on generating policy impact from IE evidence may be getting in the way of building a relationship with southern members, who are not all ready for such immediate results. It may be better to offer a wider range of services relating to quality assurance and other forms of technical assistance, building awareness and interest before seeking specific IE initiatives. A programme focussing on a few countries each year would help the national evaluation and policy communities to understand better what IE, and 3ie, have to offer.

34. Only a small group of bilateral donors is actively engaged with 3ie, and it is disappointing that there is only one multilateral member. There is a tendency for these donor members to treat 3ie as a service provider on specific programmes, rather than as the supplier of a public good for the benefit of the wider development community, weakening 3ie's core funding and strategic focus. There is a particular question about non-member donor agencies which use 3ie services and take 3ie grants.

35. 3ie's Associate Membership is described as "*a community of development experts committed to improving lives through impact evaluation*." There are 102 organisations on the current list: 12 from Asia, five from Africa, four from Latin America, and 81 from OECD countries. A number of the benefits offered to AMs are openly available on the 3ie website, which does not have a closed members' area. As with the Institutional Members, there is little to suggest to 3ie treats its associates as special. The Evaluation Team does not recommend that this should change. If, as proposed, its quality assurance role is reinforced, it might even be useful to reduce the associate membership to a core group of organisations which are willing to submit their work to 3ie for quality checking.

36. 3ie is already working to expand its membership base. However, 3ie's vision of an organisation which is 'donor-funded, southern driven' is ambitious. It will not be realised without a targeted effort, and a budget to match. The Evaluation Report identifies a need for a strategic programme and dedicated resources to support it:

- To expand the membership of Southern members, and start to build a relationship with the community of evaluation agencies in each member country,
- To remind donor members, including INGOs, of the importance of 3ie's public good function, and persuade them to continue to support a strategic, core-funded programme to deliver that public good.

3.3 The Board of Commissioners

37. In his valedictory remarks at the April 2012 Board meeting, the retiring Chairman described how 3ie Board and management had both started with little experience in this form of governance and accountability. Apart from inexperience, there were disagreements over strategy and tactics, not just between the Board and the Executive but also within the Board itself. The Evaluation Team believes that the Board has, in fact, worked rather better than many realise; certainly since a concerted effort to address institutional issues in mid-2011. 3ie has made very important progress, with limited resources. Some strain was inevitable during the start-up of such a challenging programme.

38. Until the recent creation of three new Deputy Director positions, the Executive has lacked the resources to address all the issues the Board raised. This may have led to Commissioners being more involved than would be normal, especially in their work on sub-committees. From now on, it should be possible to move to a clearer division of labour between the Board and the Executive.

3.4 Accountability in practice

39. 3ie management present a full report to the Annual Members conference, and there is discussion of strategic issues and some operational matters as well. However, the Membership is not strong enough or cohesive enough to hold 3ie to account robustly, with oversight of implementation and the final word on strategy. This is not perhaps its role. At the 2012 Members Conference a concern was also expressed about the narrowness of the representation on the new Board of Commissioners. This would not matter if there was a strong Southern voice on the Members Conference, and if the Members Conference were in a stronger position to hold the Board to account. That is difficult to arrange. With only one meeting in the early part of the year, the Conference can review the previous year's report but it cannot have much input into the next year's strategy. As to voice, the fourth meeting in April 2012 was the first one at which southern representatives were not substantially outnumbered. Even so, only four LMICs were represented, compared with nine northern organisations. The decision by 3ie's major funders to seek representation on the Board was a further step away from the vision of an organisation accountable only to the Membership as a whole.

40. 3ie and the Board have made a substantial effort to identify a better membership model. However, unless it is decided that it is not practical to include a broad southern membership, the solution to this issue cannot come from a different model. It can only come from a substantial increase in the membership, especially southern membership. Developing a membership strategy 2014-2017 would offer a useful opportunity to look at the role of the Membership Conference, in particular its role as a business meeting at which the Membership holds the Board and 3ie accountable.

4 STAKEHOLDER CONSULTATION

41. As part of the Process Evaluation, three groups of 3ie stakeholders were consulted through on-line surveys:
    - **Contract Negotiators (CN – 22 Responses).** Contract officers on grants awarded under the first three open windows, seeking views on the 3ie grant management process, including financial.
    - **Principal Investigators (PI – 31 Responses**). This sought the lead PI's views on the technical management of the grants, the implementation experience and policy uptake.
    - **Stakeholders (510 Responses).** This survey cast a wide net to capture the views of four categories of stakeholder: developing country policymakers, donor organisations, implementing agents and researchers/evaluation practitioners.

42. The surveys painted a consistent picture of 3ie as an organisation. Contract negotiators and PIs are supportive of 3ie though it does not seem that the organisation is essential to them for funding or as a technical resource. There is high awareness of 3ie's work among wider stakeholders with SRs and WPs rated highly.

43. There was a degree of consensus between the different stakeholder groups over the value of IE. All groups gave a higher priority to measuring outcomes and monitoring delivery than they did to impact evaluation, valuing the more rigorous experimental approach to impact evaluation the least. Even donor representatives gave the lowest priority to rigorous experimental IEs. Practitioners were fairly evenly balanced between viewing IEs as being essential to policymaking and being a luxury policymakers cannot afford.

44. The survey showed that as an organisation, 3ie has won wide acceptance and created an authoritative voice on impact evaluation. But there are still many who see 3ie as an advocate for rigorous, quantitative trials. Its crucial message about Theory-Based Evaluation and the need for a

balance of qualitative and quantitative methods does not seem to have been as widely received. Policymakers, implementers and donors all put experimental IE at the bottom of their list of priorities and a sizeable minority of practitioners are not convinced the case has been made for experimental IE at all. A clear strategic objective for 3ie would be to bridge this gap: demonstrating that rigorous experimental research is a necessary part of impact evaluation , and that Theory-Based Evaluation is the way to make it so.

5 PROCESS EVALUATION

45. The Evaluation Report considers six 3ie processes or groups of processes:

1. The grant management process
2. Policy and advocacy processes
3. Advancement and IE Services
4. Planning, Budgeting, Monitoring and Reporting
5. Financial Processes
6. Management Processes

To aid the assessment of the economy of those processes, it first presents an analysis of 2011 expenditures across different activities and cost categories.

5.1 3ie Expenditure by activity and category

46. The table presents 3ie expenditure for 2011. Every year is different, but data is not available for earlier years. However, 2011 seems to have been reasonably typical. With one significant exception, it matches the budget for 2012 quite closely.[2] Apart from grants, personnel costs – salaries (including benefits), and consultancy – are the major elements in 3ie expenditure: representing over 50% of non-grant expenditures. The GDN fee and travel are also significant items.

Table 2: 3ie Expenditure by Activity and Category

	Component						
	1	2	3	4	5	Total	%
Grants	9,078,848	101,796	-	-	-	9,180,644	67.1
Conference	-	293	228,026	240,412	1,503	470,234	3.4
Sub-total	9,078,848	102,089	228,026	240,412	1,503	9,650,878	70.5
Non-Grant							-
- Salaries, etc	243,646	307,392	392,370	167,176	395,953	1,506,536	11.0
- Consultancy	168,601	142,518	96,900	74,132	188,500	670,651	4.9
- Travel	88,785	33,341	80,871	47,542	186,533	437,072	3.2
- GDN Fee	694,895	-	86,862	86,862	-	868,619	6.3
- Etc	42,588	0	51,146	187	459,014	552,935	4.0
Sub-total	1,238,515	483,251	708,149	375,899	1,230,001	4,035,813	29.5
Total	10,317,363	585,339	936,175	616,311	1,231,504	13,686,692	100.0
% of Total	75.4	4.3	6.8	4.5	9.0	100.0	
% of Non-Grant	30.69	11.97	17.55	9.31	30.48	100.0	

47. The analysis shows significant shifts in the strategic allocation of 3ie's resources relative to that proposed in the Strategy 2010-2013: notably a substantial and continuing increase in the allocation to Component 3 and, to a lesser extent, Component 4.

48. 29.5% of 3ie's total 2011 adjusted expenditure was non-grant. However, it is the ratio for the two grant-giving components – 1 and 2 – which is more relevant: 16.4% and 87.2% respectively. The high figure for Component 2 (Systematic Reviews) highlights that this component's justification

[2] Adjustments made to provide the most accurate picture are explained in the main report.

must be as a standalone activity, not through grant making. Component 1, on the other hand, stands or falls by its efficiency as a grant-making operation. As such, a 16% ratio must be considered high. Two-thirds of that goes on just two items: the GDN fee and the overhead cost of 3ie's Institutional Capacity expenditure, Component 5. Without those, the non-grant ratio for comes down to 5%.[3]

5.2 The grant management process

49. 3ie's Programme Office (PO) manages the grant process in accordance with a Grant Procedures Manual issued in 2011. GDN's role as the formal contract holder is now limited to approving budgets and managing grant payments. Final approval of grants comes from the Executive Director, after the Board's endorsement of the proposed list of awards and budget. Technical review is carried out by 3ie staff, supported by a panel of External Reviewers (ER) and External Project Advisers (EPA). Approval for disbursements and letters of variation come from the Executive Director, or his deputy.

50. After a grant has been awarded, applicants are asked to respond to a set of Comments for Response (C4r) and Budget Comments for Response (B-C4r), prior to signature of the final grant agreement. Each agreement includes a full schedule of disbursements (tranches) and deliverables, against which those tranches will be paid. As well as regular progress reports, technical deliverables include draft questionnaires, survey plans, datasets, interim analyses etc. All deliverables are checked by the 3ie Evaluation Office and the EPA before a tranche is paid. Payments are also cleared by the 3ie and GDN Finance Office staff.

51. The Evaluation Report presents timelines for all 3ie's grant windows measured against four milestones: Launch, Award, First Tranche payment (i.e. contract signature on completion of negotiations), and Completion. The analysis shows it took five or six months from the launch of the major open windows to complete selection and award the grant. To complete the Conditions and Approval and Contract stages and pay the first tranche took a further six to nine months: a total of between 11 months (OW2) and 14 months (OW3) to complete the whole process. These are long timelines, especially for the Conditions & Approval/Contract stage: between three and six months longer than the target of eight months set in the 3ie Grant Procedures manual.

5.2.1 Grant Selection

52. From 70 applications for its first Open Window (OW1), interest in 3ie RfPs has increased by nearly 10 times. OW4 attracted 634 applications. To manage this volume, an Expression of Interest stage was introduced from OW3. Even so, 3ie reviewed 204 full proposals for that window. Coordinating a thorough technical review of this many proposals is a major exercise, especially since 3ie has sought to work to a high standard of objectivity and independence, on principles similar to those of the US National Research Council and the UK's ESRC. Full proposals are scored by at least four reviewers, usually one internal and three external: a long list of grants then goes to a Review Panel. The scale of the task is shown by the fact that 94 grants went to the OW2 review panel.

53. The same basic scoring framework has been used throughout, with criteria for technical quality, staff experience, developing country participation, policy relevance and policy impact. The weighting given to each criterion has been changed several times, mainly in order to adjust the balance between technical quality and policy relevance. In most cases it would seem that the technical scores are more influential. For the most recent Social Protection Thematic Window, cost was scored for the first time and a new distinction between the internal and external validity

[3] Note: for this analysis the institutional overhead of Component 5 and the GDN fee have been allocated across the active components. (See the main report for an explanation.)

of the technical proposal gave greater weight to the theory of change and the use of mixed methods. With some variations, the selection of SRs is similar.

54. For external reviewers and peer advisers, 3ie draws on a master list of 139 specialists, some of them very eminent indeed. US universities which have made experimental IE a focus of their work are strongly represented. In practice, however, 3ie has thrown the reviewer net quite widely. The one exception perhaps concerns SPTW, where a relatively narrow group of grantees overlaps with the panel of external peer advisers.

5.2.2 The Grant Management System

55. In April 2010, 3ie contracted the development of an on-line Grant Management System (GMS), with an ambitious specification. Every aspect of the process was to be managed on-line. Since then the contract has been revised two times, the current price of $432,000 is more than four times the original bid and it seems that the final delivery date is still some way off. The front end of the system is operational - registering and managing grant applications and the application review process, but the management of grants after award, has still not migrated to the GMS. If there are any changes at all in the way 3ie wants to launch a new window, the contractor has to write a special application pack at a cost of between $10,000 and $30,000 per window. While not specialists, the Evaluation Team have some experience managing IT projects. Their report registers concerns about the substantial drift in both cost and timeline for the GMS, suggesting a need for capable specialist advice and noting that IT systems have the greatest potential when dealing with high volumes of repetitive transactions: not really the case with 3ie's grant management function.

5.2.3 Comparison with Other Grant Making Bodies

56. The Evaluation Team were asked "to carry out a comparative study with similar grant making bodies, for comparison in terms of review processes, policy engagement and so on". The team reports that 3ie has worked to an unusually high standard: in its selection processes and, especially, in its management of Deliverables and Disbursements. 3ie also gives unusually high emphasis to accountability. While not on the scale of larger organisations, 3ie's guidance to grant applicants is comprehensive, and at least as accessible as that provided by, for example, the US National Science Foundation.

57. 3ie faces a bigger policy engagement challenge than research funding organisations working in the developed world. DFID's Research Programme Consortia provide a benchmark in this respect. On a quick review one RPC's website, EdQual, does not show a focus on policy influence equal to 3ie's efforts.

58. Benchmarking grant management costs requires care. Funds which disburse large sums at infrequent intervals to capable grantees can reduce the management overhead to below 5% of the total spend. If there are large numbers of small grants to grantees who require capacity building and steady support, the management cost rises sharply. DFID has prepared a comparison of large developed world research organisations, showing administration percentages between 2% and 7% of the total budget. These compare with 5% for 3ie's grant management function, without central overhead and the GDN fee; 16% with those additional costs included. However, the organisations in the DFID comparison are very large funders of first-world research: dealing at scale with long-established organisations, and in a long-established framework. The UK's ESRC has a £250 million budget: barely 1/10th goes to grants between £200,000 and £2 million, similar to those awarded by 3ie. Any firm conclusion on performance would need detailed comparison of how ESRC and 3ie deal with this type of grant.

5.2.4 Assessment

59. The evaluation team has found a clear and solid record of 3ie's grant management. Robust processes are in place for the management of frequent grant windows with 500 or more applications and up to 30 awards totalling $15 million per round. Grant selection has been rigorous. As far as is practical without field supervision, grantees are monitored closely and payments tied to acceptable delivery. By varying the scoring framework and review processes, 3ie has sought to find the right balance between technical rigour and policy relevance in its grant selection. More recently, it has also sought to widen technical assessment to include the quality of the theory of change and the use of mixed methods. Cost is now also part of the scoring.

60. At around a year from RfP to signature of the grant agreement, processes are slow. There is a need to ensure that the timeline set out in 3ie's Grant Procedures Manual is met. However, the evaluation team emphasises that this should not come at the expense of the standards being applied. Options for streamlining the process might include making a harder cut at the EoI stage. Accepting only three times as many EoIs as 3ie has funds available for would save both 3ie and many grantees wasted time. The second change would be to require a fuller final proposal, requiring applicants set out their ToC and their context analysis in more detail. This should save time during the Comments for Review process. It would also strengthen the selection process.

61. The Evaluation Team are concerned that the integrated on-line Grant Management System is proving to be an unnecessary distraction. At the current contract value, 3ie could have doubled the staff of the Programme Office for 10 years with the same money. It is recommended that the business case for the GMS be reviewed on a zero-budget basis.

5.3 Policy and advocacy processes

62. The Policy Advocacy and Communications division leads on 3ie's Strategy Components 3 & 4. It also plays a significant part in grant processes. Its work is informed by the 2010 Learning and Advocacy Strategy, and the 2011 Knowledge Management Strategy.

5.3.1 External Communication and Publications

63. The website is 3ie's main dissemination platform: as a repository of electronic versions of 3ie publications, as an inventory of 3ie projects and grant windows, and as the channel for grant funding windows, training opportunities, events and job announcements. It also provides access to the expert roster and the impact evaluation and systematic review databases. Recently revamped, the website is now more than a simple clearing house of information. It now aims to direct users to content which is targeted to their needs. A capable editorial team keeps the website fresh, as well as supporting a quarterly e-newsletter read by as many as 2,000. The same team supports the production of 3ie's publications: editing and proofing Working Papers, Systematic Reviews, and Policy Briefs.

64. Communication and publication depends on a ready supply of content, whether from IEs and SRs or from 3ie's own Working Papers and other technical documents. To date it has been difficult to develop a strategic overall approach, but this will become easier as more IEs are completed.

65. With Taylor & Francis, 3ie have published the Journal of Development Effectiveness since 2009. The publisher reports that it has been extraordinarily successful: quickly winning a very high ISI factor of 0.92. It fills a clear impact evaluation niche, with a particular emphasis on reporting negative or inconclusive results, and on policy impact. It is reported that the Journal was downloaded approximately 4,000 times in 2010. The journal has always been published to page and budget, it has a good spread of authorship from across the globe, and it has benefitted from 3ie's co-ordinated communications strategy across multiple platforms.

66. The communications team also supports the publication of 3ie's RfP's and other funding opportunities through an extensive list of 17 listservs, 40 portals, 22 evaluation networks, 8 development economics networks and 24 other communications channels.

5.3.2 Seminars and conferences

67. The Policy Advocacy and Communications team leads the major effort to organise international conferences, with the support of the GDN Travel Office, and local partner organisations. Both the Cairo and Cuernavaca conferences went well, apart from some issues over visas and travel. Participants at the latter showed a good level of appreciation for the conference in general, though there were comments about the need to more effectively bridge the gap between IE and policy and about the difficulties in applying some of the 'number-heavy' approaches to small scale community-based programmes. The London, Delhi and Washington seminar series are managed locally with the topics dependent on who is available to speak. Over 20 demand generation workshops have been held, mainly in developing countries. They have been well received, though their longer-term effect in terms of building a culture of evidence is not clear.

5.3.3 Policy, Advocacy and Communications in the Grants Process

68. Through communications strategies, and latterly Policy Influence Plans, 3ie has sought to integrate advocacy and communications into the grants process. After award, grantees are asked to submit a Policy Impact Plan for review. This identifies policy stakeholders who are asked to confirm their support for the research, both initially and at certain stages during the study. Progress on the PIPs is not a condition of tranche payment, but it is monitored, with a focus on the outcomes: e.g actions arising from a workshop. To support this process, three two-day Policy Influence Clinics have been run, with a fourth planned. Approximately 20 grantees and policy counterparts have attended each for a mix of lectures, networking and hands-on advice. It is not possible to judge how well this relatively new part of the process is working. Later in the report it is suggested some grantees may be finding it burdensome, even counterproductive.

5.3.4 Assessment

69. From the beginning, 3ie has taken a strategic approach to advocacy and communication. The small team has developed two strategy documents, and worked through them systematically, with a clear differentiation between the policy advocacy and communications. The team has made effective use of outside help to develop its strategies and support grantees. The website and e-newsletter set 3ie up as a knowledge broker for impact evaluation, and communications seem to be well received, with increasing download statistics. The publisher rates the Journal of Development Effectiveness highly. However, the subscription rates risk putting evidence generated with public funds behind a paywall, and so out of reach to many, especially researchers and policymakers in LMICs. As such it could be questioned whether it will contribute to a culture of evidence within the 3ie target audience.

70. 3ie has made a major effort working with grantees to improve the way they communicate their research results, engage with policymakers and seek policy impact. Most grantees have found 3ie's communications on this to be prompt and generally helpful. 3ie's policy and advocacy strategy papers outlined techniques to improve impact and these have been followed consistently. What those strategy papers did not discuss was the higher strategic question of whether an emphasis on policy impact is relevant to an organisation like 3ie. The implications of this are discussed later in the evaluation report, but that does not detract from recognition of the high quality of the current process.

5.4 Advancement and IE services

71. 3ie's Advancement and Impact Evaluation Services unit is based in Washington. Advancement encompasses support to 3ie members and fundraising activities. There are five Impact Evaluation Services making up 3ie's strategy Component 4: Developing the capacity to promote, use and undertake Impact Evaluations. The first four are resources for professional evaluation practitioners: an Expert Roster of IE specialists; an Impact Evaluation Database; a planned Registry of new Impact Evaluations; and a Replication programme, to encourage re-analysis and re-interpretation of IE datasets.

    3ie's fifth IE service provides quality assurance services to implementing agencies and governments, including the coordination and management of donor-funded Impact Evaluation and Systematic Review programmes. Important examples include a joint AusAid-DFID-3ie call for Systematic Reviews and quality assurance and grant management support to a very large programme of impact evaluations of HIV Combination Prevention.

72. 3ie staff also carry out a range of other Policy Advocacy and Quality Assurance activities. In the first quarter of 2012 these included: technical workshops and seminars; presentations and training for donor staff; participation in small donor discussion panels such as Bellagio; and attendance at larger events, such as the African Evaluation Association conference in Accra. There was also a Demand Generation Workshop, in the Philippines. The Executive Director makes a considerable personal commitment to this part of 3ie's work.

73. Strategic aspects of 3ie's advancement work are discussed later in the report, in the sections dealing with membership. Of the IE services, stakeholders are appreciative of the IE databases, less so of the expert roster, but this not because of any faults in the process. The Registry and Replication services are still in development. Donor agencies are making ever greater use of 3ie's Quality Assurance services, a clear recognition of their value. However, the evaluation report highlights elsewhere a risk that this may turn 3ie into a service provider, at the expense of its core public good purpose: promoting the generation and use of better IE evidence.

5.5 Planning, budgeting, monitoring and reporting

74. 3ie's Strategy 2010-2013 sets out a table of specific objectives under each of its five components, with four-year annual targets against measurable indicators. From 2011 onwards, 3ie has presented an annual Workplan including a budget allocation by activities. The activities are mapped against the strategy objectives and indicators. It has taken some time, and Board discussion, to reach this framework.

75. The Evaluation Report reviews the latest budget and workplan in some detail. A number of weaknesses are noted, starting with the fact that no comparative figures are presented for budgeted and actual expenditure in the prior years and an unhelpful presentation which combines high-level, diagrammatic presentation of the allocation of resources across strategic components with a detailed budget allocation across 31 activities and 51 sub-activities.

76. There are five more important weaknesses. First is the lack of any discussion of the availability of funds to support the budget, and the critical distinction between activities supported from unrestricted core funds and those dependent on earmarked programme funding. Second, no justification is presented for the proposed allocation between the different strategic activities - OW, PW, TW and SR funding – or between non-grant and grant expenditure. Some activities appear to roll grant and non-grant costs into one sum. Third, the only schedule of activities relates to the grant programmes. That apart, the workplan has no time dimension. Fourth, a 10% contingency is included ($1.46 million), justified on the basis that 3ie is a young organisation '*operating in a very active and dynamic field*'. This is bad practice. Lastly, there are number of new activities and budget shifts which the Board might have found it useful to consider more

specifically. Examples include a substantial increase in funding to support the membership base, and $407,000 for Policy Monitoring activities. Two expenditures are effectively capital investments with longer-term cost implications, better presented to the Board as separate projects, with a supporting Business Case justifying the commitment.

77. 3ie has published two Annual Reports, for 2009 and 2010-11. Although produced to a high standard, these seem to be designed for a wider public audience; not as an integral part of a Planning/Budgeting/Reporting framework providing management information and enabling oversight by the Board and the membership. Quarterly reporting has been irregular, with limited financial information and progress only measured against a mix of high-level output indicators, secondary indicators of stakeholder satisfaction, and purely routine items. Effective management reporting requires a rather different set of what can be termed operational indicators. These should be defined in the annual Workplan, not in the higher-level logframe. To do this, the Workplan must include a schedule of events and deliverables. In other words, it needs to be a time-bound plan, not just a list of activities and costs.

78. The 2011 financial report presented a more detailed analysis of variances against budget and commentary on the reasons. However, it still lacked a clear statement of resource availability and analysis of how under- and over-spends related to delivery against time-bound milestones. 3ie revenues were presented in terms of total grant commitments, without a time dimension. The Evaluation Report highlights the need for a cash flow analysis showing future revenues against future commitments, and separating funding available for 3ie's core activities from those dedicated to specific programmes.

5.6 Assessment

79. The Workplan and Budget are, by implication at least, the instruments by which the 3ie Board, on behalf of the membership, delegates authority for the executive to implement an agreed programme. Reporting against the agreed plan and budget should be the Board's primary instrument of oversight and control.[4] 3ie does not yet have an integrated process of planning, budgeting and reporting which is fully effective for these purposes.

80. There is no suggestion here that any major issues have arisen because of this. Until the appointment of a Finance Director in late 2011, it would have been difficult for the small and very active management team to have done more. With four new deputy directors appointed, there should now be capacity to strengthen the process substantially. The foundations for this are largely in place, in terms of financial control and time recording. From 2013 it should be possible to report 3ie's progress reported against Workplan and Budget, with variances between plan and budget reported at least six monthly.

81. The Evaluation Report makes one broader point. 3ie reports have tended to emphasise the positive without providing a great deal of detail. This approach risks raising expectations too high. The Evaluation Team heard of a certain disappointment among important stakeholders, even on the Board, that 3ie has yet to produce a large body of completed evaluations. Clearer reporting would have shown that 2012 was always the earliest date at which 3ie could be expected to deliver a substantial body of new evidence.

5.7 Financial processes

82. 3ie's financial transactions fall into four categories: grant expenditure; staff costs; professional services; travel and procurement. As processes, it is easiest to consider grant expenditures

[4] This is not stated in 3ie's Accounting Policies and Procedures Manual, which emphasises the reporting requirements of individual 3ie donors.

separately. Without in any way attempting an audit, the Evaluation Team have sought to highlight points where there may be scope to improve value for money.

5.7.1 Grant Expenditures

83. With GDN support, 3ie check each applicant's budget, seeking any necessary detail and amendments prior to signature of the grant. Typically, 40% of the grant is paid on signature, with the remainder spread over the later tranches and a final payment of 10% on acceptance of the final report. With each set of deliverables, grantees report expenditures in a standard 3ie template against six categories: personnel, travel, survey costs, equipment, other direct costs and indirect costs. 3ie's Direct Costs Policy and Indirect Cost Guideline set out the following:

    1. Personnel Costs – Gross salary, excluding overheads/fringe payments.
    2. Travel – At economy or premium economy over 9 hours, per diems at less than UN rates.
    3. Survey costs – Supported by analysis of the staff effort.
    4. Office expenses – to reflect the usage by the 3ie study.
    5. Equipment – Only below $2,500 per item.
    6. Indirect Costs – Capped 10% of direct costs for government agencies, not-for-profits and universities; 15% for NGOs and IOs; cascading to sub-grantees, 2% to the head grantee.

84. These are reasonable guidelines. The Evaluation Report reviews nine OW grants to assess how they have been applied in practice. The overall conclusion is that 3ie-GDN has endeavoured to apply the guidelines, but the results have been quite variable. Some grantees have responded to a request for more data on survey costs with a thorough and convincing breakdown. Others appear to have felt it unnecessary to demonstrate Value for Money and it is not clear that 3ie has been able to push the issue. Time and space do not allow for more than a few examples:

    OW 2.192 – 3ie-GDN queried a survey cost of $18 per 45 minute interview. The grantee provided a clear and complete statement in response. This is good practice at work.

    OW 1.43 – Co-funded with Government of Mexico, 3ie's contribution of $200,000 was to fund a follow up survey against a completed baseline. However, over half was used to cover losses on the baseline resulting from the Asian flu outbreak. It is not clear why 3ie should have covered this risk. GDN-3ie sought a breakdown of the high survey budget, but did not challenge it further.

    OW 3.1229 – This application was from a Mexican government institute. GDN-3ie asked for the 17.5% overhead rate to be reduced to 10% and accepted 15%. 45 days of conference trips was queried but stood. A budget of $130,000 and 478 staff days seems high for re-analysis of existing data and a small, semi-structured survey of 129 interviews.

85. The Evaluation Report notes some more general points. First, co-funding creates uncertainty, especially where 3ie is a relatively minor partner. Secondly, there were a number of grants which involved the agency implementing the intervention, including one where the agency was the main recipient of the grant funding. The cost of PI time is quite variable. GDN-3ie checks on VfM seem to be focussed on the budget. Checks at later stages seem less robust. Provided the deliverables are accepted, tranches are paid without any comparison of actual expenditure against budget. On a number of grants, the IE as implemented is significantly different from the original proposal, but the budget was not adjusted.

86. The report emphasises that time only allowed a rapid read of the material. Nevertheless, the Evaluation Team believe there is a case for more robust and consistent application of 3ie's cost policies, with a default position that they are not negotiable. Funder support may be necessary where this results in a clash with the financial policies of other institutions. Closer attention to Value for Money, <u>as delivered</u>, may be particularly important; with the option of early termination if the possibility of an outcome commensurate with the cost becomes too small. Matching tranche payments more closely to the expenditure profile should also be considered. The stronger

finance team established in late 2011 should have the capacity to support this more robust approach. However, the EO and EPAs' judgement on whether the effort is justified by the scope of work is fundamental.

5.7.2 Non-grant Expenditures

87. Apart from the GDN fee, 3ie's principal non-grant expenditures fall into three categories: salaries/benefits (37% of non-grant expenditure); consultancies (17%); and travel (11%).

88. 3ie's staff costs are largely determined by those of the two hosting organisations: GDN and LSHTM. The Evaluation Team have not reviewed salary levels or attempted to benchmark them against comparable rates. They were advised that GDN rates are not very attractive in the Indian market. There is perhaps an impression that the opposite is true in London and Washington. In the former, salaries at the lower level seem high.

89. 3ie's consultancy budget is split between support to the grant management process and more ad hoc requirements. For the first, External Reviewers and External Peer Advisers receive a standard fee of $200 per review. For a simple review of high quality applications, a reviewer might complete three, at most four reviews a day. However, it is the Team's experience that it is difficult to keep this rate up, and the reviewer's comments show that they have put a substantial effort into many of the reviews. Overall, this service has given 3ie good value.

90. The Evaluation Team have not analysed 3ie's use of ad-hoc consultants. This category ranges from specialists contracted to produce technical IE and SR guidelines to legal advisers hired by the Board of Commissioners to assist with governance issues. It is possible that the capital costs of IT development are also included in this category. However, if the Component 1 cost, mainly for IE review, is excluded, the annual spend under this category in 2011 was just over $500,000: high enough to justify a review of the way this work is commissioned.

91. Most 3ie travel is managed by the GDN travel office. This is an important operation. 3ie staff travel frequently and events such as Board meetings and the bi-annual conferences increase the load substantially. The 3ie GDN MoU indicates that the use of the GDN travel office is included in the hosting fee, but it may be that events are charged separately.

5.7.3 Hosting Agreements

92. At $868,000 in 2011 The GDN hosting fee was 3ie's second largest non-grant cost, after salaries and benefits. The base fee of $600,000 covers rent for the E. Wing of the GDN Office in Delhi, four GDN staff dedicated to 3ie operations, IT support, and the use of the Travel Office and other services. 3ie pays the capital cost of furniture and equipment through an expansion payment for each staff post.

93. Two elements of the GDN agreement merit attention: the rent, and the percentage management fee. The rent of $200,000 is for a large, well-appointed modern office in a mid to high level area of Delhi. The Team is not qualified to judge how it compares with market rates, or how important it is that 3ie has a prestige address. 3ie pays 7% of grant disbursements as a management fee, falling to 6% above $6 million. In Q1 2012, 3ie made 81 payments (about 6 per week). Given that 3ie is paying a full cost for office rent, etc., it is hard to see how this generates a burden on the GDN overhead equivalent to 7% of the grants disbursed.

94. The hosting arrangement with the London School of Hygiene and Tropical medicine was put out to competition in the UK and the School's bid was the lowest. A London location was an additional advantage. The relatively high salaries, especially at the Research Assistant level, have already

been mentioned.[5] In other respects the arrangement seems quite cost-effective, especially if LSHTM is carrying the cost of pension and insurance cover for the staff.

5.7.4 Procurement

95. Grant management aside, procurement is not a major function for 3ie. The Accounting Manual set out a simple three-tiered set of procedures which should be adequate for most requirements. However, 3ie has awarded three substantial contracts: Policy Influence Monitoring (PIM) ($400,000 - ODI), The Grant Management System ($400,000 - Praxis), and, the IE Registry ($790,000 - RAND). The first of these was a competitive and well-run tendering process. Six proposals were reviewed by a four-person panel with two external members against a detailed RfP. The other two are both IT development contracts. The Evaluation Report ascribes significant slippage in time and cost on the first of them to a lack of the specialist capability needed to manage this kind of procurement robustly. The second contract with the RAND corporation has only just started. On the material made available, the report expresses a concern that similar problems may occur with this much larger investment.

5.8 Management processes and style

96. 3ie has been a very reflective organisation. In mid-2011 it commissioned three reviews of its internal operations, including a 360° review of the Executive Director's performance. In a small organisation, his part has been fundamental: working to win recognition for 3ie as a centre of excellence for Impact Evaluation, at the same time as developing the systems to manage a substantial programme of grants and supporting activities. The Evaluation Report presents an overview of the conclusions of the 2011 review, with limited comments on how the concerns identified are being addressed.

5.8.1 Institutional Reviews

97. The first Rapid Institutional Review describes 3ie as an organisation "of exceptional delivery and performance, with good progress toward the Initiative's strategic objectives". Two concerns are noted:

- Communication. 3ie's message of success to internal stakeholders was "being diluted and undermined by anxiety and needless concern."
- Over-commitment. "The Initiative is overcommitted and overextended and expectations are unrealistic and unreasonable"; rapid organisational growth was "being handled on an ad hoc basis without a systematic understanding of the capacity in place …."

To address these concerns the report recommended that 3ie adopt a performance management system, measuring performance against realistic targets, and establish its own HR function.

5.8.2 The Leadership and Organizational Performance Review

98. This review had three parts: a 360° review of the ED, a Culture and Climate Survey of 3ie staff, and a survey of 3ie Members. The first showed a high degree of support for the ED from most groups and on most aspects of his job. His direct reports are particularly supportive. Lower scores from Executive Committee and the wider Board suggest some tension in this relationship. The ED's self-ratings demonstrated good awareness of his own weaknesses, and coincide in most respects with those of others. Given what has been reported elsewhere, it is perhaps revealing that the one exception is the Direct Reports' lower rating on Communication and Delegation.

[5] This is assuming the stated salaries in the MoU do not include a margin to LSHTM

99. The Climate and Culture Survey was is similarly supportive. 80% or more of the staff have a positive view of working with 3ie. Once again the lowest scores (50-67%) were in the area of performance management. Communication was another area in which there more 'agrees' than 'strong agrees' with the positive view. A low score on the statement 'I see myself working in this organization 3 years from now' was disappointing: just 46% in agreement.

100. Five donors participated in the member survey: Gates, DFID, Hewlett, AusAid and DANIDA. On a four point agree-disagree scale, 3ie was scored positively but not massively so across questions on strategic direction, the quality of grant proposals, the management of member funds, 3ie's achievements and the ED's performance. The lowest scores had to do with internal capacity and management systems, fund management and the ED's efforts to diversify the donor base. Real progress was noted, with impressive results and an established position in the IE community. But the members felt 3ie was too small to have an impact, and shared the concern about overstretch and the need to decide its priority objectives.

101. The Evaluation Team were particularly interested in the members' comments on the quality of 3ie grants, a point to be picked up in detail in Section 6.3. Although the methods seem rigorous and the PIs are recognised experts, they felt there was not enough information to gauge the ***quality*** of proposals, a top priority during 3ie's start-up. They were uncertain that the diversity of 3ie's portfolio was the most effective way to build an aggregate base of evidence in the strategic areas.

5.8.3 Assessment

102. The concentration of reviews in mid-2011 suggests a degree of concern at that time. The reviews themselves show a consistent picture, one that is well summed in the words of the donor members interviewed for the Leadership and Organisational Performance Review. They recognised 3ie's remarkable progress during its first phase and the hard work of the staff, the ED in particular: 'a major achievement'. However, it was now time to focus on developing 3ie's strategic direction and to shift from the intensity of a "start-up" to the more structured approach of a mature international organization.

103. A number of steps have been taken towards that transition. Deputy Directors have now been appointed for Finance and Administration and for the operational units. The foundations of a stronger financial management system are largely in place. However, as discussed in Section 3, 3ie is still some way short of the robust planning, budgeting and report structure needed to maintain focus on 'standing and important objectives', and to communicate those objectives clearly to staff as a basis for managing their performance.

104. It is stressed that this should not mean that 3ie refuses to respond new needs and opportunities as they arise. As a resource for all its members, that response is an important part of its function. What is required is a clear strategy to manage the balance between strategic objectives and new opportunities, and to ensure that such opportunities are not undertaken unless adequate spare capacity is available.

6 RELEVANCE AND EFFECTIVENESS (CH 6)

105. Chapter 6 of the Evaluation Report considers two different aspects of 3ie's Relevance and Effectiveness: the Theory of Change, with its the underlying logical framework; and, the Potential Impact of 3ie grants, as the best immediate measure of the effectiveness of 3ie's efforts to generate rigorous, policy relevant evidence.

6.1 The theory of change

106. 3ie's current Theory of Change has two levels. At the first level, advocacy for impact evaluation leads to more IEs and synthetic reviews being conducted, thence to the dissemination of the evidence generated, and to increased capacity to undertake further research. At the higher level this influences policy, through the better use of evidence, and leads to better development outcomes. Feedback loops improve 3ie's recognition and generate more funding for IEs, and so on. Key impact measures are the number of 3ie-funded studies completed and published, the number of policymakers influenced by the evidence produced, and the number of cases in which development policy design or implementation is demonstrably affected by that evidence. Developing country involvement is central to the theory, made explicit in impact indicators for the participation of LMIC evaluators and policymakers.

107. The report discusses each step of the Theory of Change in detail, looking at both the horizontal and vertical relationships and the line of argument between them. It ends by proposing the revised Theory of Change presented in Figure 2 below.

6.1.1 Assessment

108. The following assumptions are implicit in the ToC:

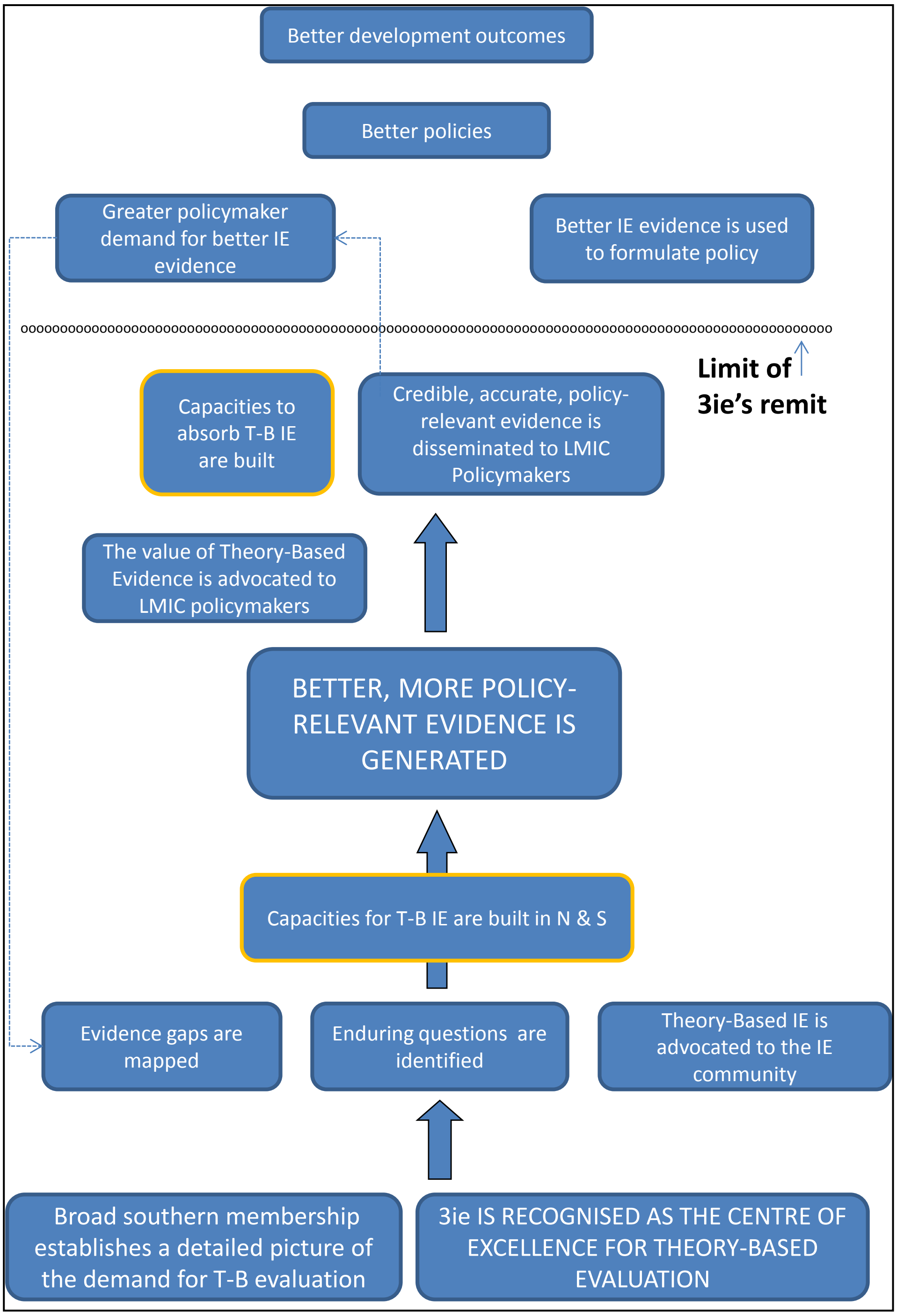
- That the overriding need is for a greater quantity of rigorous evidence,
- That the quality dimension can be taken as given: that more IEs will automatically generate that greater quantity. The possibility that the need is for IEs which are more rigorous, as opposed to more IEs, is not considered,
- That policy relevance can also be taken as a given: that the research community will generate enough policy relevant proposals and it will only be necessary to select the most policy relevant,
- That policy influence requires a major effort which must be built into research from day one.

109. The Evaluation Team believe there are problems with each of these assumptions. The stakeholder survey shows that many policymakers remain to be convinced that rigorous IEs should be an important part of their policy processes. The next section of the report sets out reason to suggest that the quality dimension cannot be taken as given. Researchers are facing significant difficulties translating experimental/quasi experimental techniques to the development context. Policy relevance is also quite variable. There seems to be something of an inverse relationship, such that the most policy relevant programmes are least amenable to experimental IE techniques, and *vice versa*.

110. The Evaluation Team present an alternative Theory of Change starting from two propositions. First, that there is an urgent need for greater rigour in development IEs, and that rigour cannot come solely from the application of statistical technique. It is 3ie's Theory-Based approach which is the true gold standard for impact evaluation. The second proposition is that policy relevance can only be determined in developing countries. 3ie has a clear role to play in working with a broad southern membership to identify where that relevance is. By combining these two strands, 3ie will deliver one single purpose: <u>the generation of better, more policy relevant evidence.</u> The proposed new ToC places a limit on 3ie's remit, seeking impact solely through the quality and relevance of the evidence, without additional efforts to influence policy directly.

111. Figure 2 showing the proposed alternative Theory of Change is presented overleaf.

Figure 2: Proposed revised 3ie Theory of Change



6.2 The logical framework

112. The Evaluation Team have been asked to review the Logical Framework, in particular the output indicators. The review of the Theory of Change would suggest a radical revision, but this should be led by the 3ie team and involve all major stakeholders. The Evaluation Report only presents a

revision of the current logframe, to strengthen it as a tool for monitoring 3ie's progress.[6] The following changes are proposed:

1. GOAL. A reversal of the phrase 'better use of evidence' to 'the use of better evidence' strengthens the logic. 3ie's mandate is to support the production of more, better evidence; not to ensure that evidence is used in a better way.
2. PURPOSE: The one major revision is to introduce a Purpose statement: "*IE evidence influences development policy and implementation".* This change indicates that it is 3ie's Purpose to deliver IE evidence which influences policy and implementation. That this influence will lead to better development outcomes is a fundamental assumption of the Theory of Change. But it is not something that can be directly attributed to 3ie. Indicators of policy influence fit the logic better at this new Purpose level: e.g. 'programme designs changed by IE evidence', 'institutions influenced by IE evidence'. Two critical assumptions are moved up: the political limits to reform, and the effectiveness of IE evidence. One new Purpose indicator is proposed to bring the concept of Enduring Questions into the logframe, on the basis that answers to these questions will have the greatest policy influence.
3. Four assumptions have been struck out. This is because they are either internal to the programme, i.e. something which 3ie should influence directly, or they are a killer assumptions: in other words, if they fail, then 3ie should never have been started.
4. A set of more concrete Output/Objective indicators is proposed, with less dependence on secondary, subjective measures of 3ie's performance such as stakeholder satisfaction. Indicators of output quality are suggested: e.g the number of IEs published in refereed journals and citation scores. Greater weight is given to summary indicators of operational progress: e.g grant applications received, and grants awarded.
5. A particular effort has been made to develop more specific indicators on 3ie's Policy Window and LMIC government initiatives and simplify the link to Objective 1 through the involvement of LMIC researchers.
6. Two key indicators of institutional strength are added: the level of core funding available to support 3ie's work, and the Membership. Two existing indicators are hardened as: 'Number of Board actions completed/uncompleted', and 'Workplan objectives delivered.'

6.3 The potential impact of 3ie grants

113. Recent debates over the evaluation of the Millennium Villages and de-worming programmes have highlighted how any evidence on development is subject to robust challenge, and rightly so. Controversy over the Cochrane Collaboration's review of de-worming has shown that this just as true for systematic reviews as it is for impact evaluations. As a body with a specific mandate for standard setting, quality and credibility are particularly important for 3ie.

114. With 9 IEs and 9 SRs completed, the evidence 3ie has generated is only just beginning to enter the policy process. To assess its potential for policy impact, the Evaluation Report reviews the two most critical determinants of that impact: the quality and policy relevance of the research it has funded. It is important to emphasise that the small team had very limited time for this review. Inevitably it reflects the personal experience and judgements of the specialists involved. Readers are recommended to refer to the Evaluation Report (Section 6.3) and, if possible, to review the source material so they can weigh the evidence for themselves. That said, the Evaluation Team is confident that the review presents a reasoned assessment of the quality of 3ie's portfolio, sufficient at least to highlight a number of key issues and provide a response to a concern expressed by 3ie donor members that *"Although the methods seem rigorous and the PIs are*

[6] The logframe shown in 3ie's Gates Revised Proposal 5 July 2010 is taken as the starting point.

recognised experts, they felt there was not enough information to gauge the quality of proposals, a top priority during 3ie's start-up. They were uncertain that the diversity of 3ie's portfolio was the most effective way to build an aggregate base of evidence in the strategic areas."

115. 3ie have produced a robust set of practice papers and guidelines underlining how rigour is not just about the statistical design, summed up in the ED's presentation statement that "Measurement is not Evaluation". The Evaluation Team believes 3ie's emphasis on Theory-Based evaluation is of the utmost importance; to the extent that they feel 3ie itself may even be underselling it. They have specifically sought to judge 3ie grants against this broader standard of rigour: looking at the quality of the theory of change, and the detail with which the causal chain is described, as well as the statistical design. Two other important aspects are considered: the grantees' understanding of the context and the quality of study implementation, in particular questionnaire design and application.

6.3.1 Make-up of the 3ie Grant Portfolio

116. The IE portfolio is reasonably balanced between Sub Saharan Africa and S. Asia, with smaller numbers in E. Asia and S. America. The majority of grants are in five sectors: health, social protection, agriculture, education and private sector development. 3ie-funded studies cover three categories of development intervention:

- Major on-going programmes, such as Ethiopia's Poverty Safety Net Programme, and W. Bengal's metering of electrically powered tubewells.
- The introduction of a recommended policy to new areas, such as Intermittent Screening and Treatment for malaria among Kenyan schoolchildren.
- Relatively untried interventions such as cook stoves in Ghana, mobile banking in Sri Lanka and a new pension product for micro-finance customers in India.

It is easier to organise a clean and statistically effective design for studies in the second and especially the third category. Conversely, immediate policy relevance is more likely with the first.

117. A number of 3ie studies test issues of behavioural economics, in particular the way beneficiaries and service providers respond to different incentive structures. For many PIs 'Nudge Theory' seems to be an essential part of the toolkit. These issues are amenable to controlled-trial forms of evaluation, but this focus limits attention to a rather short length of the causal chain.

6.3.2 Specialist Review of Four Sectors

118. The Evaluation Report presents an independent specialist's assessment of each of the five main sectors. The specialists were specifically chosen as leading development practitioners, not as evaluation specialists. Against an overview of enduring, policy relevant issues in the sector they were asked to assess: i) the theory of change set out in each 3ie-funded study; ii) its understanding of the context; iii) its potential contribution to current knowledge; and, iv) the likelihood of generating useful conclusions for policy makers.

119. The Evaluation Report presents the specialists' ratings against the four assessment points on a five point scale. Taking a score of at least three across all four points as the measure, 17 out of 63 were rated as genuinely strong (27%). At the other end of the scale, 13 grants did not score more than 2 on any point (21%). Of these, six were in Finance/PSD, three in Health, three in Social Protection and two in Agriculture.

120. The specialists were also asked to review the sectoral portfolio as whole: its coherence; its potential impact; its relevance to known policy questions; and particular strengths and weaknesses. Space does not allow the detailed conclusions to be presented and the following summary can only be incomplete. Readers are again recommended to refer to the Evaluation

Report to weigh the evidence for themselves. (As 3ie's first thematic window, Social Protection is discussed separately below.)

Coherence – The reviewers found the grants they looked at to be collections of separate research studies, rather than coherent research portfolios. There were some areas of focus – ECD in education, food fortification in health, weather and market risk management in agriculture. In agriculture the studies have the potential to contribute to important debates. In health, on the other hand, the reviewer felt the concentration on fortification was disproportionate. In Education, having five studies in countries with large populations is a clear strength and the IE and SR portfolios are complementary. It would have been useful to use the SR findings in grant selection.

Relevance – Almost all the reviewers identified major sectoral issues which were not covered in the portfolio: "making markets work for the poor" in PSD; scaling up and health systems strengthening in Health; survival, retention, quality and equity in Education; and equity and gender in agriculture. A number of reviewers felt that grants reflected relatively narrow academic research interests. One mentioned the dominance of economics departments, as opposed to institutions with recognised strengths in the sector.

Policy Impact – studies linked to existing policy changes or programmes do have the potential for real policy impact. The agriculture portfolio is strongest in this respect. In Health, about a third of grants were rated to be of both high quality and relevance. Another third were of high quality but studies of "a small element of a sub-topic, nothing approaching an enduring issue." In several Finance studies, the main beneficiary seems to be a private sector entity, with no consideration for any potential for scaling up or poverty reduction. The Agriculture reviewer noted a similar concern. The Education reviewer emphasised that what policy makers need most is policy options with clear analysis of cost effectiveness, to support their decisions on how best to invest each dollar. In that sector proposals varied, but there did not seem to be a clear definition of the concepts of cost-effectiveness. The two final reports available did not deliver what the proposal promised on cost-effectiveness.

Theory of Change – Few grants state an explicit Theory of Change, 'beyond brief references to causal links.' This leaves it unclear how the findings are expected to inform policy makers. Links between intermediate and final outcomes are not systematically explored. The short duration of many studies is an important limitation. Some only look at short term impacts; others seek to measure impacts on livelihoods, education and health from data gathered over just two years. As one reviewer puts it, "Short term studies can only realistically measure first stage impacts: adoption, yields, gross margins, etc. Apparent success on these indicators may mask the wider policy implications." Many grant proposals indicate that mixed-methods will be used but few show clearly how this will be used to complement the quantitative findings, and there is little evidence of it being done in the completed reports.

Context – One reviewer described 'contextual naivety', in particular over the quality of available data, calling into question 'the carefully designed edifice of the RCT'. Even if researchers do demonstrate an understanding of country context, this does not translate into a clear summary of policy context, questions of political will, the role of the media and other social factors.

6.3.3 Statistical Review of Completed IEs

121. As part of the output review, Triple Line invited an internationally recognised statistician and biosocial scientist with wide experience in developing countries to review the final reports of nine studies. After confirming that 3ie's guidelines provide clear procedures for planning and executing rigorous IEs, the specialist undertook a detailed review of the reports. A simple assessment scale was adopted - if the report had been submitted to an international peer reviewed journal would it be publishable without revision, with some minor revisions, with major revisions, or not at all?

122. The results were quite disappointing. Only one study report was found to be ‘well written and sound’. Many of the others were strongly criticised for lack of rigour, unjustified policy recommendations, and poor presentation. Only two of the nine papers were considered to be definitely publishable. While the Evaluation Team believes the reviewers’ analysis to be authoritative, it must be recorded that one study he judged as ‘a poor report which if submitted to a journal would either be rejected or would require substantial revisions,’ has in fact been published by the BMJ. Asked to reconsider his judgement the reviewer largely maintained the view that “This trial has ‘succeeded’ not through the incentives provided by the researchers to reduce anaemia but through pre-existing incentives to improve academic performance.”[7]

123. This example underlines how even considered judgements over such an apparently clear cut matter as the validity of an RCT can prove to be quite subjective. This in turn reinforces the importance of 3ie’s role as a setter of IE standards and provider of IE quality assurance.

6.3.4 IE Implementation Review

124. The Evaluation Report reviewed 19 grants across the three open windows in more depth. The sample was chosen to cover a balance across the windows and thematic sectors, and to include both the low and high end of reviewer scores during selection. The first step was to review what happened at each of the five stages in the grant process: application, selection, contracting, deliverables & disbursement, and final report. This analysis revealed two things: the intensity of the quality assurance effort 3ie has put in, and the fact that this effort was clearly necessary. Although it cannot be expected that there will not be significant issues on some grants, the number of significant issues in a small sample of grants is cause for concern.

125. To try to throw more light on this, the Evaluation Team leader personally reviewed eight grants from Open Windows 1 to 3. Subject to there being sufficient data to allow a detailed review, the grants were selected to give greatest weight to the larger OW2 and to be evenly spread across the scoring 3ie reviewers gave the grants when selected. The Evaluation Report presents a summary of his findings. Given the time available and the volume of material to review, these responses can only give an impression, and they are subject to checking. The discussion is too lengthy to include in this summary report. Once again readers are recommended to read the relevant section in the Evaluation Report to weigh the evidence for themselves (pp. 126 to 130). They are urged to do this because the conclusions are not encouraging.

126. All eight grants show significant problems of implementation and analysis. For a number of them the study as delivered is, or clearly will be, significantly different from what was proposed. Poor preparatory work, or reconnaissance, has meant that the proposed design cannot be implemented. This is most important where its springs from the failure of a critical assumption. Whether it is very low uptake of condoms in Zambia or very low demand for pre-school places in Mexico, these failed assumptions are highly policy relevant; more policy relevant, the reviewer argues, than the treatment, the actual intervention being evaluated. On other studies it is clear that the treatment was untested, even not fully designed, prior to the proposal to roll it out for the evaluation study. Examples include mobile banking without a functioning transfer system; and a pension scheme which is not clearly defined in the grant proposal, and which the provider is not yet ready to support. One study is of a Sustainable Land Management treatment which is, on the evidence presented, untested. It appears to be assumed that the treatment will offer farmers no financial benefit: supported by the most unlikely statement that “[agricultural] adoption itself has no objective monetary value.”

127. In some cases, poor reconnaissance has led to major statistical re-designs and changes in method, undermining the rigour which is IE’s primary objective: ‘augmented’ samples; imputed values for 20% of the survey sample, with the basis of the imputation unexplained; and,

[7] The full review can be found in Annex 2.

replacement rates of 16% and 20% in one survey. Where available, survey questionnaires are found to be excessively long: 76 pages with a blood test, in one example, and enumerators given a minimum time to complete of 2 hours 35 minutes in another. The latter questionnaire includes questions which are likely to be impossible to answer sensibly, or analyse; for example asking poor Indian women to respond to behavioural economics tests on the preferred cash outcome from a coin toss. In another, the data shows that pumping hours, the critical parameter on which the final report depends, has been calculated from what can only be farmers' most approximate recall estimates of days pumped and hours per day. Multiplying the two figures can only multiply the inaccuracy of the original estimates.

128. Statements in progress reports that "individuals were more exposed to financial products than assumed", or that there were unexpected problems with the public extension service, reveal that grantees have not reviewed available material on either the treatment or the context. In most cases, this is extensive and provides a more than adequate base for a robust, theory-based evaluation design; a design which takes account of known issues in the sector concerned. A study of electricity metering in the W. Bengal market for tubewell irrigation is a clear example. It takes no account of two fundamentals, soil type and depth to groundwater; it makes only the briefest reference to extensive literature on how this very competitive market adjusts returns to land, labour and water; it does not analyse how relative price changes have affected the market, particular how metering has affected the price of irrigation; and, it makes no mention of the policy issue which metering is intended to address: the reliability and volume of electricity supplies. Instead the study is modelled as a narrow, abstract question of behavioural economics: a two-person game between the pump owner and the farmer; with no evidence presented on how that relationship works in practice.

129. It is important to stress that this review has only been made possible by 3ie's rigorous management of grantees, in particular the requirement that they submit concrete deliverables: draft questionnaires, preliminary data, etc. The record also shows that 3ie's internal and external reviewers, the External Peer Advisers in particular, raised many of the issues set out in the Evaluation Report with the grantees. What they do not seem to have been able to do is make the points with sufficient force to get them addressed effectively. This is understandable, given the high reputation of the PIs and research institutions involved in the grants.

6.3.5 3ie's Systematic Reviews

130. Although a few grants have been very delayed, 3ie has managed the SR grants robustly to deliver a body of capable reports on water and sanitation issues, mother and child health, genital mutilation, HIV prevention in women, social cohesion and early child development. This is reinforced by SRs on water and sanitation and farmer field schools completed internally.

131. Almost all of these are technically strong reports, applying rigorous selection to ensure that only studies which controlled for endogeneity and self-selection bias were reviewed. Despite searching very widely and considering up to 20,000 studies for inclusion, as few as six were found which met those standards in each sector. Even among these relatively small numbers, the variation between intervention approaches leaves room for uncertainty about how much of the measured impact is attributable to the basic intervention and how much to the implementation approach. In short, the availability and quality of source material is one of the biggest constraints, with many reaching the weak policy conclusion that more research is needed. Four out of six SRs published in 3ie livery effectively present that conclusion.

132. 3ie faces opposing risks. One is that in the attempt to present a 'policy friendly front-end' to its systematic reviews, it will slip into promoting policies for which the evidence is weak. The opposite risk is that a repeated finding that more research is needed leads policy makers to conclude that useful evidence will never come. This is another area in which the Evaluation Team believes that Theory-Based Evaluation and mixed-methods are fundamental, because they will

make it possible to draw policy conclusions while the quantitative evidence is still weak. The Social Cohesion SR, for example, does a neat analysis of how intervention characteristics may affect project outcomes, identifying how “interventions have not been carried out in accordance with [CDD] theory.” This qualitative conclusion is both policy relevant and more important, than the finding that the statistical evidence is weak and more studies are needed. That said, systematic reviewing has an important negative function: challenging IE practitioners to do better. The current debate over de-worming stimulated by Cochrane Collaboration review is a good example. A conclusion that more research is needed may not be policy relevant. A conclusion that better research is needed is highly policy relevant, especially so given the results of the IE review presented above.

6.4 Enduring questions and themes

133. From the 3WEL paper onwards it was expected that 3ie would work to identify and find answers for enduring questions about how to improve social and economic development programmes. How to do this has itself become an enduring question for 3ie, one that has yet to be clearly resolved, with continued debate about the balance between responsive and strategic approaches, in particular between Open and Thematic Windows for IE grants. 3WEL identified ‘Enduring Questions’ as a specific activity for the international initiative on IE, and this was carried through into the DFID Project Memorandum and the MOU with BMGF.

134. 3ie made a significant effort to meet these expectations, culminating in an Enduring Questions Workshop held in Kuwait in early 2009 which considered responses from a wide range of stakeholders. The workshop identified a range of possible themes which were taken to the Board, but it is understood that the Board could not reach a consensus on this thematic or sectoral approach. Some felt that the Open Window approach would provide sufficient strategic focus, as long as the selection took account of the potential for policy relevance and policy impact. This led to some tension with 3ie’s principal funders. With the possible exception of Early Childhood Development, the sector specialist reviews for this evaluation report do not suggest that grants made under the Open Windows have come together to offer a solid body of evidence on particular themes or enduring questions

135. There has been debate about the balance between open calls for IE grant applications and focused, thematic windows. 3ie launched its first thematic window in 2011: the Social Protection Thematic Window. The Evaluation Team have been asked to assess SPTW separately, to review the advantages and disadvantages of the two modes of operation.

136. An Expert Advisory Group was formed for SPTW and a Scoping Report presented an inventory of social protection programmes and mapped knowledge gaps and research questions. It suggested that SPTW focus on social transfer programmes, looking at enduring issues under two headings: impacts and design. For the first, research questions include the contribution of social transfers to economic growth, purchasing power and labour market effects; and how social transfers affect education and health outcomes and traditional support practices, gender aspects in particular. In the area of design, key debates include targeting, the use of conditions, the concept of graduation; and how to link social transfers to complementary interventions. A key ideological debate is whether social protection should be inclusive or be targeted at the poor, with the aim that they should graduate from poverty. To handle this polarised debate governments need rigorous evidence so they can account for their social protection programmes effectively. This also means that contextual factors such as political will and capability are critical.

137. The Triple Line specialist reviewer found the Scoping Report provided a good conceptual framework for SPTW. However, this opportunity was missed, in that the final portfolio of grants did not respond to that framework. Of the 10 SPTW studies:

- A number looked at impact, but only two studies on Ethiopia's Productive Safety Nets Programme are likely to discern long term change. For most, the analysis of cost-effectiveness and cost benefit is weak; an area the scoping report correctly identified as a considerable concern to governments and development partners.
- No studies focus on political economy or institutional analysis, or how different components of social protection can be integrated into a comprehensive system. All ten study a single programme in isolation, more than half of them small-scale, even pilot programmes, with limited likelihood of gaining the political support for scaling up.
- Design issues are not emphasised. One or two studies compare alternative delivery channels. Some consider the issue of exit/graduation and the interface with complementary programmes; but none compare targeting options or look at the significance of human capital conditionality: two key areas of contention in the current debate. There were none of the recommended "cross-country or cross-regional studies".
- With only three studies of national-scale social protection programmes, and six of small-scale, many essentially donor-driven, the reviewer felt the portfolio is unbalanced.
- Four of 10 grants are to IFPRI, two of them looking at the same project, and two to the University of North Carolina, studying the same type of project in the same way, and for the same purpose, in neighbouring countries.[8] This could be seen as an economical way to build a solid body of evidence on critical issues, but it will also limit the potential richness of the research.

138. The reviewer concludes that 3ie's social protection portfolio is useful, but not dramatically so. The focus on small-scale pilot programmes ignores the realities of social protection: that proven impact is no guarantee of political support or cost-effectiveness at a larger scale. In the worst case, the result may be proof of positive impact, but a pilot which ends leaving its temporary beneficiaries no better protected than they were before.

139. 3ie's scoring criteria, and application and grant review formats for SPTW made no reference to specific research questions or to the areas identified in the Scoping Report. A tighter specification and a requirement that grantees justify their proposal against the areas identified in the Scoping Study might have resulted in a stronger portfolio.

7 CONCLUSIONS AND RECOMMENDATIONS (Ch. 7)

140. The Evaluation Team have had extensive access to 3ie's records and open consultation with a broad range of stakeholders. They have reviewed of a reasonable sample of 3ie outputs in detail and feel able to present these conclusions with confidence. It is important, nevertheless, to stress that no evaluation can capture every aspect. The conclusions are presented under two headings: Relevance and Effectiveness, and Process Issues, distinguishing between relatively minor issues and those which have more importance for 3ie's operation and strategic direction. Two final sections present four major strategic challenges, and an outline for a Strategy to address them.

141. It is important to state the very substantial progress 3ie has made in its first three years:

- An enthusiastic, highly capable team has been established.
- A large portfolio of IE and SR grants has been commissioned through rigorous, objective processes.

[8] The impression of concentration is reinforced by the fact that an IFPRI staff member is EPA on the 2 N. Carolina grants and another is Lead PI on one grant and EPA on another.

- Strong systems are in place to monitor the quality of grant implementation.
- An authoritative set of guidelines and working papers present 3ie's vision of best practice in both Impact Evaluation and Systematic Review.
- 3ie is recognized across a wide, diverse group of evaluation stakeholders.
- After early difficulties, an effective working relationship has been established between the Board of Commissioners and 3ie management.

These are very substantial achievements for a start-up organisation. It is difficult to see how 3ie could have achieved more.

7.1 Relevance and effectiveness

142. 3ie has been highly effective in winning recognition for its technical capability and the quality of its services. While many others have contributed, it has played an important part in making the case for funding to fill the Evaluation Gap. Through its own grants it is making an important contribution to filling that gap.

143. 3ie's mandate can be clarified and this will reinforce its relevance. 3ie has won recognition as a funder/manager of IEs; as an advocate for IE; as a builder of capacity for IE; and, as a standard setter in IE. The Evaluation Report argues that it is the last which defines 3ie's most important role. No other organisation has given the same importance to setting out what constitutes rigorous, policy relevant impact evaluation. The case for IE has been made, but there is still a need for advocacy to support 3ie's standard setting role. That is, more specifically, advocacy for Theory-Based Evaluation, and for a much stronger acceptance that when it comes to international development rigour can only come from a balanced combination of quantitative and qualitative methods. 3ie is the only organisation playing this role and it has a clearer vision on these matters than any other. However, it has not yet been effective in communicating that clarity. There remain many who only see 3ie as an organisation which funds IEs, and promotes statistical rigour and experimental technique.

144. 3ie has faced the classic dilemmas of grant fund management. If selection standards are set too high, there may be no grants worth funding. The effort required to extract an acceptable result from a grant, once awarded, can become so great that it might have been better to do the work in-house. The Evaluation Team believe 3ie has made every reasonable effort to select credible grant proposals, and to hold grantees to a high standard. In some cases it has clearly gone the extra mile. Despite those efforts, the overall standard of the evidence produced is disappointing. There is no reason to believe that other organisations have done better.

145. It was inevitable that a rapid increase in IE funding would stretch available capacity. Some of the better known practitioners seem to be over-committed. There are still many lessons to be learnt about how to apply rigorous experimental techniques in circumstances faced by poor people in developing countries. This implementation challenge has been underestimated and few studies base their theory of change on a good understanding of existing development literature, which is both substantial and closely argued.

146. The Evaluation Team believe 3ie is uniquely equipped to address these weaknesses. It has a demonstrated commitment to theory-based evaluation, and to a balance of quantitative and qualitative methods. Its specialist staff and reviewers have a very strong understanding of experimental design and statistical technique. If this were combined with the same calibre of theoretical and practical knowledge of the sectors 3ie works in, it would stand as a true centre of excellence for Impact Evaluation. 3ie's proposed Replication Service would be an important tool. If donors and governments were to require say 10% of all IEs they fund to be replicated independently by 3ie, the quality of IE studies might improve dramatically.

147. The success 3ie and others have had in making the case for increased IE funding has obscured a key point in the original Evaluation Gap discussion: that Impact Evaluation is a public good best delivered by an independent organisation representing all stakeholders. 3ie's membership structure was intended to achieve this. However, there is a tension between this and the fact that most funds for evaluation are tied to particular sectors. Convincing donors to provide core funding for real public goods, setting standards and managing high quality studies of enduring questions, will be difficult. Nevertheless, the Team believes that unless 3ie makes the effort to maintain its higher, public-good function it risks becoming a service provider which, however effective, must tailor its strategy to the availability of dedicated, mostly sectoral funding.

7.1.1 Building Capacity

148. 3ie was not expected to have a major capacity building role. It has strengthened its requirements on the involvement of developing country researchers, but it does not require a grant budget for capacity building. Overall, 3ie's impact on capacity is unlikely to be large, but it was not expected to be.

7.1.2 Influencing Policy and Use of Evidence

149. 3ie has made a major effort to integrate measures to influence policy, improve the use of evidence and generate policy impact into its grants. On balance, the Evaluation Report concludes that this may have led in the past to an over-emphasis on policy influence, as opposed to policy relevance. Requirements for the latter are relatively clear: that the evidence should be valid, cover the issue in question, confirm or challenge current hypotheses, and explain the underlying relationships. Policy influence, on the other hand, is conditional on the interactions between different groups and the quality of the policy process, not aspects which 3ie can affect. It is more important that 3ie is a trusted source of independent evidence on what works and why, on what questions are most relevant to policy, and on the standards needed for a valid answer to those questions. The Evaluation Report concludes that 3ie should focus on good research communications, and rein back efforts to demonstrate policy influence at this stage.

7.2 Process issues

150. Two minor process issues merit attention. First is the procurement and management of major non-grant contracts, IT contracts in particular. Secondly, as for any start-up organisation, performance management and HR are important issues. On balance, the Evaluation Team do not agree with an earlier consultant that 3ie needs a dedicated HR function. It is enough that a member of the new, stronger senior management be given specific responsibility for this role. At present, it is the Delhi-based Deputy Director for Finance and Administration. It might be helpful if one of the more senior technical managers at least shares this task with him, to ensure 3ie's large technical team is both represented and effectively managed: perhaps the Deputy ED.

151. There are three more major process issues. In order of importance they are: (i) cost; (ii) planning, budgeting and reporting; and (iii) membership.

i) The Evaluation Team are wary of a conclusion that 3ie's administrative costs are high relative to other grant funding agencies. That said, the GDN hosting arrangement seems poor value for money and there is a case that 3ie's institutional costs are high. The distributed structure across Delhi, London and Washington imposes additional costs. Although it is difficult to see an effective alternative, there may be a case for moving some functions from London and Washington to Delhi.

ii) Until recently 3ie's financial control and management reporting were weak. The elements of a robust planning, budgeting and reporting system are in place. It is important that they are now brought together, at the latest for the 2013 budget, but ideally with stronger quarterly reports for the second half of 2012.

iii) Since the Founding Document set out an inclusive vision of 3ie's Membership, the inability to realise that vision has been 3ie's own Enduring Question. Although principally a strategic issue, there are process elements. Better reporting to the membership as a group, and a concerted effort to communicate regularly with the membership may be part of the solution.

7.3 Strategic challenges

152. 3ie faces four major challenges: Policy, Quality, Relevance, and Membership.

(i) **Policy -** For 3ie's evidence to have an impact on development policy and practice it must achieve three things: quality, and the credibility which goes with it, relevance and influence. After careful consideration the Evaluation Team believe 3ie has in the past placed too much emphasis on the last, in particular on the mechanisms of influence**.**

(ii) **Quality -** This leads to the second point, quality. Quality issues are not unique to development. In developed countries, both medical and social science research is strongly contested, with doubts raised about the quality of research and the incentive structures behind it. It is no criticism of 3ie to say that the quality of the grants it has funded is disappointing. On the contrary, 3ie has worked at least as hard as any other organisation to set a high standard for its grantees and challenge them to meet that standard. The Evaluation Team believes that it has worked harder, and it is in a unique position to do so. Its own technical base is strong and clearly set out in a set of robust guidelines, although its theory-based approach merits greater emphasis. Unusually, 3ie combines work on IEs with SRs, giving an overview of the entire evidence-generation process. And, by combining a strong theoretical understanding with direct experience of grant management, it is in a unique position to see how theory translates into practice. The Evaluation Team believes that setting standards and checking that they are met should be 3ie's fundamental job. This will not be easy. An IE industry has developed which believes it is doing good work. The unresolved debate between quantitative and qualitative approaches testifies to that. To have an impact on quality, 3ie will need the concerted support of the users of impact evaluation: policymakers and donors. But quality is the true public good in impact evaluation for development. As is recognised in the *When Will We Ever Learn Paper*, which put quality as the first job for an international evaluation initiative.

(iii) **Relevance -** A strong theory of change and contextual analysis is just as critical for the policy relevance of IE studies as it is for their quality. Although debates continue, only a minority now questions the need for rigorous impact evaluation. Apart from encouraging more funding, 3ie's most important role has been to convince the evaluation community of the need to address the attribution challenge, in particular to develop experimental or quasi-experimental study designs which effectively control for selection and other biases. However there is important advocacy work still to do: to encourage researchers to develop more effective ways to tackle the attribution challenge in the diverse, shifting contexts of development, and to emphasise the importance of standards. While recognising the case for an Open Window approach to grant commissioning, the Evaluation Team believes that Thematic Windows will be more effective. The first TW may not have achieved the focus set out in the scoping study, but experience should improve the process. The tendency towards a menu approach, where funders select questions within broad themes, may weaken the focus. This is another argument for suggesting that 3ie should run its own TWs, on themes analysed by 3ie staff and approved by the Board and, ideally, the Membership.

(iv) **Membership -** Although process changes may strengthen 3ie's relations with the Membership, an increase in the number of Members is the only real solution to this issue.

And that will require a targeted and funded effort to achieve. For southern members, the Evaluation Team suggest a broader approach, based more on advocacy and services than on generating demand for IEs per se. This will not be easy: to support it, the Team recommends that 3ie begins drafting a Strategy 2014-2017, which emphasises the public good nature of 3ie's work. With the Board's support, it could then initiate a campaign to enlist donor members to provide the core funding needed to implement it.

7.4 Outline for a revised strategy

153. The Evaluation Report presents an outline for a revised 3ie strategy, based on the conclusions set out above and linked to the proposed new Theory of Change presented in Section 6.1, and to the simple rewording of 3ie's goal from 'better use of evidence' to 'the use of better evidence'. Instead of encouraging a larger volume of impact evaluation evidence, 3ie's main purpose should be to raise the quality and policy relevance of all impact evaluation evidence. In other words, it should aim to become the Centre of Excellence for Impact Evaluation

154. The six elements of such a strategy would be:

- 3ie becomes the **standard-setting body for impact evaluation** and works to show how far the IE evidence being produced meets the standards it sets. As well as setting high standards for its own grantees, it would offer to certify studies funded elsewhere to the same standard. A sample of grants would be replicated, or independently re-analysed, as a matter of course. If funders back this approach, there is scope for a significant improvement in IE standards. Ideally, it will also reduce current debates around the validity of different IE efforts, something which is particularly unhelpful for policymakers.
- 3ie makes a concerted effort to communicate the **message of theory-based evaluation**. It would also add a sectoral dimension to its review processes and give greater weight to broader, non-statistical theory.
- 3ie reins back from its current focus on demonstrating policy impact and instead focuses on ensuring that the **IEs it commissions are of sufficient quality** to present to policymakers and are policy relevant.
- As part of this, 3ie will continue to develop the Policy Window and Thematic Window approaches and make TWs part of its core funding programme.
- 3ie develops a membership strategy which centres on understanding the **breadth of southern members' demands for impact evaluation evidence** and providing a range of services those members need to realise that demand.
- 3ie develops a Strategy 2014 – 2017 along the lines described, and campaigns to win core funding for the public goods it will provide under the programme.