

Final Report

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
111 Westminster Bridge Road
London
SE1 7JD, UK
Tel: +44 (0)20 7922 0300
odi@odi.org.uk
www.odi.org.uk

Table of Contents

1	Introduction.....	1
2	The Current Status of 3ie.....	2
2.1	Strategy.....	2
2.2	3ie Background.....	2
2.3	Progress to Date: Grants.....	4
2.4	Policy Advocacy and Communications.....	12
2.5	Feedback from 3ie Applicants and Others.....	20
2.6	Organisation and Staffing.....	24
2.7	Structure and Staffing.....	24
2.8	Management Systems.....	27
2.9	Funding Status.....	29
3	Governance, Direction and Accountability.....	33
3.1	Constitution and Bylaws.....	33
3.2	The Membership.....	36
3.3	The Board of Commissioners.....	39
3.4	Accountability in Practice.....	39
4	Stakeholder Consultation.....	41
4.1	Contract Negotiators.....	41
4.2	Principal Investigators.....	43
4.3	Stakeholder Survey.....	47
4.4	Conclusions.....	59
5	Process Evaluation.....	62
5.1	3ie Expenditure by Activity and Category.....	62
5.2	The Grant Management Process.....	64
5.3	Policy and Advocacy Processes.....	73
5.4	Advancement and IE Services.....	79
5.5	Planning, Budgeting, Monitoring and Reporting.....	80
5.6	Financial Processes.....	86
5.7	Management Processes and Style.....	90
6	Relevance and Effectiveness.....	94
6.1	The Theory of Change.....	94
6.2	The Logical Framework.....	101
6.3	The Potential Impact of 3ie Grants.....	105
6.4	Enduring Questions and Themes.....	121
6.5	Conclusions.....	126
7	Conclusions and Strategic Review.....	131
7.1	Process.....	131
7.2	Relevance and Effectiveness.....	132
7.3	Strategic Review.....	133
	List of People Met.....	135

Tables

Table 1: Budget Allocation	2
Table 2: Grants Awarded Since 2009.....	5
Table 3: Progress of the 3ie Grant Portfolio.....	6
Table 4: Delivery Profile of the 3ie Impact Evaluation Portfolio	6
Table 5: Financial Status of 3ie Grant Portfolio - March 2012	11
Table 6: 3ie Policy Advocacy and Communications Objectives.....	12
Table 7: PACO History - summary to date.....	14
Table 8: Completed Projects - Communications and Advocacy outputs	15
Table 9: Cuernavaca Conference Expenditure (460 participants).....	17
Table 10: 3ie Seminar Series.....	18
Table 11: Summary of 3ie Publications and Advocacy Efforts	20
Table 12: Stakeholder Surveys carried out by 3ie, Triple Line and others	21
Table 13: 3ie Payments to GDN.....	26
Table 14: Funding Status of 3ie - 2008 to 2016.....	30
Table 15: 3ie Cashflows - Actual and Projected 2008 to 2016	31
Table 16: Stakeholder Response Numbers.....	41
Table 17: Usefulness of 3ie Publications	49
Table 18: Usefulness of 3ie Web Resources.....	49
Table 19: 3ie Expenditure 2011 by Component & Category	62
Table 20: Relative Weighting of 3ie's Components.....	63
Table 21: Analysis of Overhead Cost Distribution	63
Table 22: Score Weighting for Grant Proposal Reviews - SPTW	69
Table 23: Summary of GMS progress	70
Table 24: 3ie QA, Training and Advocacy Activities	80
Table 25: 3ie Specific Objectives and Targets for Component 1.....	81
Table 26: Detailed 2012 Workplan / Budget for 3ie Component 1.....	82
Table 27: Workplan Indicators, 2009-2012	83
Table 28: Sample of SR Team Workplan	92
Table 29: Draft for a 3ie Logical Framework	103
Table 30: Breakdown of the 3ie Portfolio of Impact Evaluations.....	107
Table 31: A Specialist Practitioner Assessment of 3ie IEs	109
Table 32: Data from OW2.082 on Hours of Water pumped from Tube Wells in W. Bengal.	119

Figures

Figure 1: Impact Evaluation Grants by Length 8

Figure 2: 3ie Associate Member Responses 23

Figure 3: 3ie Organogram 24

Figure 4: 3ie Governance Structure 34

Figure 5: 3ie's Importance to Stakeholders 50

Figure 6: Influence of Evidence in DC Policymaking 51

Figure 7: Influence of Evidence on Policy Formation? 51

Figure 8: Policymakers' M&E Preferences 52

Figure 9: Donor M&E Preferences 54

Figure 10: Policy Use of IE Evidence – Researchers and Practitioners 56

Figure 11: Which is more important timeliness or quality? - Researchers/Practitioners 57

Figure 12: Quantitative vs. Qualitative 58

Figure 13: What are RCTs worth? 59

Figure 14: The Grant Management Process 65

Figure 15: Timelines for 3ie Grant Windows 67

Figure 16: Administrative Costs of Large Research Funding Organisations 71

Figure 17: Policy Advocacy, Communications and the grant process 77

Figure 18: The 3ie Theory of Change 94

Figure 19: Mission creep in advocating for robust evidence in policy processes 98

Figure 20: Mission creep between good research communications and policy influence. 99

Figure 21: Suggested Theory of Change 100

Acronyms

3ie	International Initiative for Impact Evaluation
3WEL	When Will We Ever Learn
A&C	Advocacy and Communications
A&FC	Audit and Finance Committee
AfrEA	African Evaluation Association
AIDS	Acquired Immunodeficiency Syndrome
AIES	Advancement and Impact Evaluation Services
BC4R	Budget Conditions for Response
BMGF	Bill and Melinda Gates Foundation
BMJ	British Medical Journal
C2E	Commitment to Evaluation
C4R	Conditions for Response
CCAS	Culture Climate and Awareness Survey
CCT	Conditional Cash Transfers
CDD	Community Driven Development
CEPA	Cambridge Economic Policy Associates
CGD	Center for Global Development
CIPPEC	Centro de Implementación Políticas Públicas para la Equidad y el Crecimiento
CN	Contract Negotiators
COI	Conflict of Interest
CRD	DFID Central Research Department
DAC	Development Assistance Committee
DANIDA	Danish International Development Agency
DD	Deputy Director
DFG	German Research Foundation
DFID	Department for International Development (of the UK)
DIME	The Development Impact Evaluation Initiative
ECD	Early Childhood Development
ED	Executive Director
ELO	External Liaison Officer
EM	Evidence Matters
EO	Evaluation Office
EOI	Expression of Interest
EPA	External Program Advisor
EQ	Enduring Questions
ESRC	Economic and Social Research Council
FD	Founding Document
FFS	Farmer Field School
FTE	Full Time Equivalent
FWO	Fonds Wetenschappelijk Onderzoek
FY	Financial Year
GBP	Pound Sterling
GDN	Global Development Network
GMS	(Online) Grant Management System
EOI	Expression of Interest
HIV	Human Immunodeficiency Virus
HR	Human Resources
IDB/IADB	Inter-American Development Bank
IDCG	International Development Coordinating Group

IDRC	International Development Research Centre
IDS	Institute of Development Studies
IE	Impact Evaluation
IFMR	Institute for Financial Management and Research
IFPRI	International Food Policy Research Institute
IIED	International Impact Evaluation Department
INGO	International Non-Governmental Organisation
IO	International Organisation
IPA	Innovations for Poverty Action
IST	Intermittent Screening and Treatment
IT	Information Technology
JDE	Journal of Development Effectiveness
J-PAL	The Abdul Latif Jameel Poverty Action Lab
KM	Knowledge Management
LIDC	London International Development Centre
LIC	Lower Income Country
LOV	Letter of Variation
LMIC	Lower and Middle Income Countries
LSHTM	London School of Hygiene and Tropical Medicine
M&E	Monitoring and Evaluation
MAER	Meta-Analysis of Economics Research
MCC	Millennium Challenge Corporation
MDG	Millennium Development Goals
MG-NREGA	Mahatma Gandhi National Rural Employment Guarantee Act
MIC	Middle Income Country
MIT	Massachusetts Institute of Technology
MoU	Memorandum of Understanding
NGO	Non-Governmental Organisation
NIH	National Institutes of Health
NL	Newsletter
NONIE	Network of Networks on Impact Evaluation
NORAD	Norwegian Agency for Development Cooperation
NSERC	Natural Sciences and Engineering Research Council of Canada
NSF	National Science Foundation
NWO	The Netherlands Organisation for Scientific Research
ODI	Overseas Development Institute
OECD	Organisation for Economic Cooperation and Development
OW	Open Window
PACO	Policy, Advocacy and Communications Office
PB	Policy Brief
PI	Principal Investigators
PO	Program Office
PIM	Policy Influence Monitoring
PIP	Policy Influence Plans
PSNP	Productive Safety Net Program
PPG	Project Preparation Grant
PW	Policy Window
PWPG	Policy Window Preparation Grant
RAPID	Research And Policy In Development
RCT	Randomised Control Trials
RFP	Request For Proposals

RFQ	Request for Qualifications
RIR	RAPID Institutional Review
ROMA	RAPID Outcome Mapping Approach
RPC	Research Programme Consortia
SCF	Save the Children Fund
SLM	Sustainable Land Management
SMART	Specific, Measurable, Attainable, Relevant and Timely
SNF	Swiss National Science Foundation
SP	Social Protection
SPTW	Social Protection Thematic Window
SR	Systematic Review
TB	Tuberculosis
T-B	Theory Based
ToC	Theory of Change
ToR	Terms of Reference
TLC	Triple Line Consulting Ltd
TW	Thematic Window
US	The United States of America
VET	Vocational Education and Training
VfM	Value for Money
WASH	Water, Sanitations and Hygiene
WP	Working Paper

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 Triple Line Evaluation Team started its assignment on 16 April 2012, when it attended 3ie's 7th Board Meeting in Rome. Since then the Team has visited 3ie's headquarters in Delhi, for two weeks in May, and made two visits to the London office. The Process Evaluation has been carried out in line with 3ie's Request for Proposals, Triple Line's proposal and an Inception Report submitted after the briefing visit to the Rome Board Meeting.
2. 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. The Evaluation Team would wish to record their special gratitude to the 3ie Executive Director and his staff: for their openness and their patience with and prompt response to un-ending requests for more data, and for the way the Team were made welcome to the Delhi Headquarters. The Team must also record their appreciation for the way the Commissioners made them welcome at the Rome meetings, especially for their generosity with their time during a busy period, and for their openness in the interviews.
3. Triple Line has also carried out an e-Survey of three groups of stakeholders: **Contract Negotiators** on grants awarded by 3ie; **Principal Investigators** on those grants; and, a **broad community of policymakers**, donor officials, and evaluation practitioners: over 5,000 emailed and 500 responses. 57 key stakeholders have been interviewed individually: Board Commissioners; 3ie management and staff; policymakers in developing countries; donor evaluation units; and, people who have played a key role in the development of the Impact Evaluation movement.
4. The Evaluation Report has seven Chapters:

Chapter 2 sets out the Current Status: progress to date in terms of grants issued and completed policy advocacy and communication; and the development of 3ie's systems. It also reviews the funding situation.

Chapter 3 summarises 3ie's constitution and reviews governance and accountability structures, in particular the critical issue of Membership.

Chapter 4 presents the results of the three e-Surveys.

Chapter 5 sets out the main Process Evaluation assessing each of 3ie's major processes: grant management; policy advocacy and communications; advancement and IE services; planning, budgeting, monitoring and reporting; financial processes; and, management processes and style.
5. As part of the Process Evaluation, Triple Line was asked to make a preliminary review of 3ie's relevance and effectiveness; preliminary because the final impact cannot be known until later. To answer this requirement the Evaluation Team attempts an assessment of 3ie's Relevance and Effectiveness in **Chapter 6**. After looking at the Theory of Change and logical framework, a review of the Potential Impact of 3ie Grants forms the central part of this assessment: seeking to judge the quality and policy relevance of the evidence generated with 3ie's support.
6. **Chapter 7** summarises the overall conclusions of the Process Evaluation and what those conclusions suggest for the future direction of 3ie's strategy.

2 The Current Status of 3ie

2.1 Strategy

7. 3ie’s strategy is set out in **Improving Lives Through Impact Evaluation: A Strategy for 3ie**. An explicit Theory of Change (ToC) is presented, seeking to address the evaluation gap from both the supply side, “the insufficient production of impact evaluations”, and the demand side, “the failure to use the (IE) evidence that is available.” (Section 6 below presents a detailed analysis of the strategy and the ToC.) The strategy document also sets out a four year programme: 2010 – 2013. With a total budget of \$51.4 million, the programme has five components:

Table 1: Budget Allocation

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

8. With over 75% of the total budget, the strategic focus on commissioning IEs to generate new, policy-relevant evidence is clear. To use the strategy report’s own language, there appears to be a strong emphasis on the supply side. However, it is made clear that the aim is not just to fund studies by leading international practitioners in impact evaluation. It is also to build a commitment to, and capacity for, evidence-based policy-making through the involvement of evaluators and researchers from LMICs, and by funding demand-led evaluations proposed by government agencies and NGOs. Funding through dedicated Policy Windows to generate demand-led evaluations is a central element of Strategy Component 1.

2.2 3ie Background

9. 3ie’s origins go back to 2004, when the Center for Global Development (CGD) launched its Closing the Evaluation Gap Initiative, with support from the Hewlett and Bill and Melinda Gates Foundations. CGD coordinated an Evaluation Gap Working group and led a process of consultation and discussion, culminating in the 2006 paper *When Will We Ever Learn (3WEL)*; which can be described as the manifesto of the Impact Evaluation movement.
10. 3ie springs directly from a 3WEL recommendation that ‘some entity whether a committee, standards-based network, secretariat, or other organization—is needed as a focal point for leading’ an initiative to overcome the evaluation gap. Many of the issues which 3ie now faces spring from the vision set out in 3WEL, and how that vision has been interpreted.
11. The vision was for a membership organisation including any of “*developing country governments, development agencies, NGOs, foundations, and other public and private entities that volunteer to generate new, policy-relevant knowledge about social programs.*” (The emphasis on social programmes, rather than development programmes more generally, was explicit.)

12. 3WEL described four core functions for this new entity:
- *Establishing quality standards for rigorous evaluations.*
 - *Administering a review process for evaluation designs and studies.*
 - *Identifying priority topics.*
 - *Providing grants for impact evaluation design.*
13. Other possible, though less central functions, were suggested: managing information, dissemination, and capacity building, etc. Critically, there was no consensus over whether or not this new entity should be a grant making body, managing a pooled fund to commission impact evaluations on behalf of its members. Some felt this would just divert funds from existing efforts. Others argued that the public good nature of evaluation, and the very existence of an evaluation gap meant that a new funding channel was needed. Besides, an initiative without major funding would lack credibility.
14. Following further discussions and consultations over the next two years, a core group of ‘provisional member institutions’ came together to support the establishment of 3ie: the Bill and Melinda Gates Foundation, the Hewlett Foundation, DFID, Google.org and the Netherlands Development Co-operation agency. The initiative was formally registered as a non-profit corporation in the US state of Delaware on 14 May 2008. A Founding Document was agreed in June of that year. The member institutions had already selected an Executive Director, contracted initially through CGD.
15. The 3ie Founding Document interprets the 3WEL vision as follows:
- 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.
16. The Founding Document set out detailed provisions for the membership of 3ie and for its governance. It made it clear that 3ie would be a grant making body, but one which would concentrate 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.*” Eight principal activities were named:
1. Identify enduring questions about how to improve social and economic development programmes.
 2. Identify development programmes which offer opportunities for learning.
 3. Adopt quality standards to guide its reviews of impact evaluations (IE).
 4. Finance the design and implementation of IEs that address questions of enduring importance to policymaking.
 5. Prepare syntheses to link findings from individual IEs with broader policy questions.
 6. Advocate for the generation and use of IEs.
 7. Disseminate information about opportunities for learning, planned studies, designs, methods, and findings.

8. Promote the mutual development of capacity to implement IEs and to use evidence in policymaking

17. The Founding Document stated that it was to be ratified when 3ie had at least three members from Low and Middle Income Countries (LMIC), three members from bilateral/multilateral agencies, and three members from not-for-profit organisations. An Addendum to the document set out interim operating procedures to be followed until the formal establishment of the 3ie Board of Commissioners. The founding members named a board of six interim commissioners to manage the transition process. In the event, it took nearly a year to establish the Board of Commissioners, which held its first meeting in Paris on 17 June 2009, and took over from the interim board. At this point considerable progress had already been made: the Executive Director and a small team were contracted through GDN or on consultancy contracts; a first Open Window for impact evaluation grants had been completed with the award of 18 grants; a Systematic Review Window had awarded 7 grants, and an agreement had been signed with GDN for it to host 3ie and provide administration and support services.

2.3 Progress to Date: Grants

18. 3ie awards two principal categories of grant: for Impact Evaluations (Strategy Component 1) and for Systematic Reviews (SR – Strategy Component 2). To date, it has run the following competitive funding windows:
 - 3 Open Windows (OW) and one Policy Window (PW) for IEs
 - 3 Systematic Review (SR) windows
 - It is currently running a fourth OW, a fourth SR window, a second PW and a Thematic Window (TW) for impact evaluations on social protection.

19. As of 30 June 2012, 3ie had committed \$35.2 million in grants under the completed rounds. Of this, 54 per cent, \$19 million had been disbursed.

20. Table 2 summarises the volume of applications and grants awarded under the nine different funding windows 3ie has launched since 2009.

Table 2: Grants Awarded Since 2009

Window	Date Launched	Date of Award	No of E.o.Is	No of Applications	No of Grants
OW1	Jan-09	Jun-09	-	79	18*
OW2	Sep-09	Jan-10	-	255	30
OW3	May-10	Oct-10	366	204	22**
OW4	Jun-12	-	634	-	
SR1	Nov-08	Q1 2009	20	17	7
SR2	Aug-09	Q3 2009	-	10	2
SR3	Oct-10	Jan-11	-	24***	6****
SR4	Feb-12	May-12	-	?	?
SPTW	Oct-11	Mar-12	220	84	10****

Notes: * One OW1 grant was cancelled; ** One OW3 grant has since been cancelled; *** Joint AusAid, DFID and 3ie call. 24 of 187 applications for 3ie research questions; **** Grant agreements under process at time of writing

21. Table 3 analyses disbursement and progress with grants awarded under the completed funding rounds. Grants are paid in up to six tranches, at between six and 12 month intervals. A substantial initial tranche is paid on signature of the grant agreement. Payment for further tranches is tied to 3ie's acceptance of agreed deliverables for each one. As such, a comparison of payments made against the agree tranche due-dates gives an indicator of grant implementation progress.

Table 3: Progress of the 3ie Grant Portfolio

Round – Date of Award	Grants	Av. Grant \$'000	Av. Duration Months	Total Value \$'000	% Disbursed	Grants Completed	Grants Overdue *
IMPACT EVALUATIONS	68						
OW1 – 2009	17	249	32	4,231	69	5	1
OW2 – 2010	30	504	29	15,111	69	1	7
OW3 – 2010	21	423	31	9,011	47	0	3
POLICY GRANTS	3						
PW1 – 2011	3	220	4	660	14	1	2
SYSTEMATIC REVIEWS	15						
SR 1 – 2008	7	50	7	349	85	5	2
SR 2 – 2009	2	89	22	179	75	0	2
SR 3 – 2010	6	72	12	431	27	0	5
TOTAL	86			29,972**		12	27

Notes: * Grants overdue have tranche payments outstanding and an end date before April 2012, 2+ tranches outstanding/end date before July, 3+ tranches outstanding/end date before October.

** Small Project Preparation Grants etc. make up the remainder of commitments.

2.3.1 Open Window Grants

22. Of the \$30 million committed at March 2012, by far the largest part, \$28 million or 93%, is for grants awarded under the three open windows. OW2 had the largest number of grants and the largest average grant size. It alone represents over 50 per cent of the total.
23. Table 4 sets out the delivery profile of the current portfolio. It highlights how 2012 is the year in which 3ie expects to have its first substantial results to present. With only six grants completed out of 26 due at the time of writing, midway through the year, there remains a substantial delivery challenge to meet that expectation in the next six months.

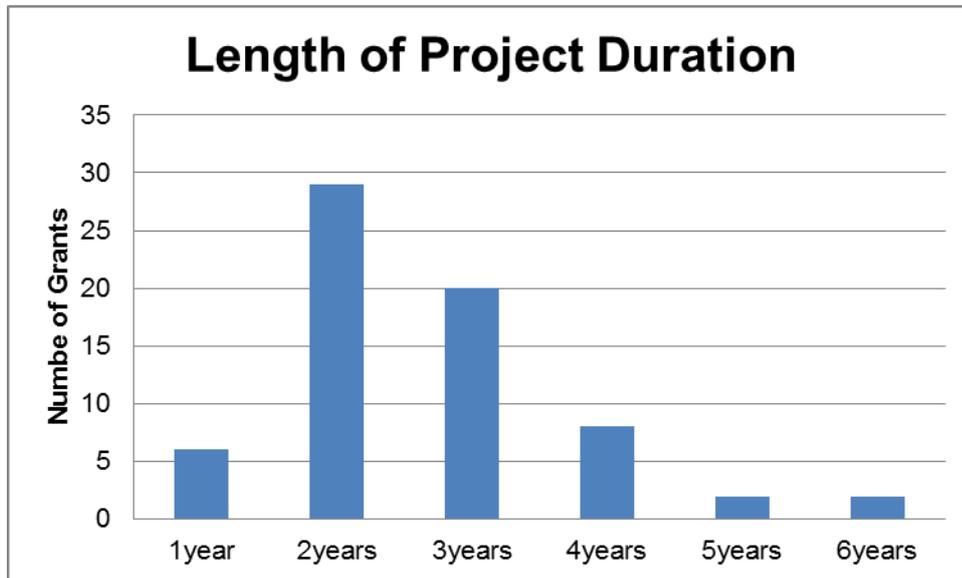
Table 4: Delivery Profile of the 3ie Impact Evaluation Portfolio

Round, Due/ Delivered	2011	2012	2013	2014	2015
OW1 grants due	2	10	2	1	2
Grants delivered	2	3	-	-	-
OW2 grants due		12	11	7	
Grants delivered		1	-	-	-
OW3 grants due		4	6	9	3*
Grants delivered		0	-	-	-
Total grants due	2	26	19	17	5
Grants delivered	2	4	-	-	-

Note: * one OW3 grant due in 2017 cancelled, (Mauritania study OW3.1217)

24. With just over three years since the first grant awards, OW1 is furthest advanced. 18 grants were awarded, but one was cancelled, as the result of a change in Government. Final reports have been accepted for six and one more is at draft final stage. In the case that a time extension or budget variation is needed, 3ie issues a Letter of Variation (LoV). Seven OW1 grants have had LoVs for time extensions, for as much as a year, and two had budget increases.
25. OW1 was launched in early 2009 and the review process was complete by June of that year. Grant agreements were scheduled to be signed and first payments made by November, and most of them were. However, a small number were not signed until mid-2010, and the first payment on one grant was not made until February 2011. It is understood that some delays were caused by difficulties in contracting between GDN and certain universities from the United States. These were resolved by having grants held in a separate institution: in most cases IPA.
26. OW2 was launched in Q4 of 2009, and most grant agreements were signed between June and August 2010. To date, one grant is complete and another is at draft final stage. Seven OW2 grants have experienced delays and nearly half, 14 grants, have had Letters of Variation. Some LoVs relate to delays in single tranche payments but there have also been 11 budget amendments and some extensions of the end dates. Some budget amendments were minor, others were substantial: in one case 30% of the original grant, over \$100,000. Some budgets were reduced. One very substantial reduction (over \$300,000), was the result of increased funding from the Gates foundation which was co-financing the study.
27. OW3 was launched in Q2 2010 and the majority of grant agreements signed in August and September 2011. To date four LoVs have been agreed including one three month extension of end date and one change of Principal Investigator (PI). Two other grants without LoVs are delayed. For one of these a tranche due in September 2011 has still not been paid.
28. OW4 launched in early 2012. 634 EOIs were submitted and 250 proposals were received by the deadline of 6 August 2012.
29. In his March 2012 report, the Executive Director advised the board that fewer studies had been completed than expected, as is indicated in Table 4 above. Among the main reasons he reported were:
 - Delays in implementing the intervention to be evaluated
 - Fieldwork taking longer than anticipated
 - Study teams take longer than expected to revise reports
 - Two grants have been cancelled because the intervention did not take place: one because of a change in government, and the other because of conflict.
30. This issue will be reviewed in more detail in when looking at the grant process. (Section 5.2) Overall, the delays in grant implementation do not seem to be greater than might be expected for a large, complex programme of studies implemented by a wide range of organisations. It is also important to recognise that two-thirds of 3ie IE grants have been awarded for studies taking two years or more, as summarised in Figure 1. Provided that the large batch of grants due to finish in 2012 are completed by the end of the year, 3ie will have largely met its targets. There is perhaps more room for concern over the length of time taken to complete the commissioning process. The span from launch of the window to awarding of the contract compares well with other organisations which have an external review process. However the time from contract award to signature of the grant agreement and payment of the first tranche can be long: over a year in the case of OW3 (though this has reduced for SPTW). The review of grant processes in section 5.2 below gives particular attention to this aspect.

Figure 1: Impact Evaluation Grants by Length



2.3.2 Policy Window Grants

31. Making grants available through Policy Windows for demand-led impact evaluations proposed by governments and NGOs is a key part of the 3ie strategy. The first policy window was launched in Q2 2010. A two-stage process was planned whereby implementing agencies nominated development programmes for evaluation and then competitive proposals were sought to carry out the evaluation. 33 applications were submitted from policy makers and scored on policy relevance and potential policy impact. Nine were selected, but two from Mexico and two from Mozambique were merged together. The addition of two proposals from OW2 considered more appropriate for the Policy Window brought the number back up to nine.
32. The second stage of the process was not successful. Eleven proposals were approved by the PW1 review panel, but problems arose in developing full proposals. Five proposals were taken through, but one (Malawi) was withdrawn because the intervention did not go ahead. In two others (Mexico), the researcher concluded that a robust IE design was not possible. As a result, only two applications were approved. One, in Uganda, was a full grant; the first tranche payment has been made and fieldwork is expected to start in 2012. In Fiji, a preliminary Project Preparation Grants (PPG) was made to facilitate the development of an acceptable full proposal, though the process needed considerable support from 3ie and a replacement international team. The report on possible evaluations is now being discussed by the Fijian government. The net result has been that no evaluations will be completed for any of the 33 programmes nominated by implementing agencies under Policy Window 1.
33. To try to overcome these difficulties, PW2 is following a different approach. Through a demand generation process, 3ie staff work with policy-makers and implementation agencies to identify programmes for evaluation. To date, Demand Generation Workshops have been held in the Philippines and S. Africa. At the end of this process, 3ie signs an MoU with the implementing agency and issues a Request for Qualifications asking researchers/evaluators to apply to carry out the evaluation(s). A Policy Window Preparation Grant is awarded to the best qualified research group. If this results in an acceptable proposal, it is followed by award of an IE grant. Currently PW2

initiatives are underway in four countries, RFQs have been launched in four others and are under discussion in yet more countries.

- **India** – PWPG was granted to design IE for an intervention to combat early marriage. Project is underway and first tranche paid.
- **China** – REAP deworming study and a study on infant supplements awarded. A further study on computer assisted learning has been rejected. MoUs with 2 Chinese government agencies for both projects.
- **Ecuador** – A PWPG awarded to prepare up to 3 proposals for Ministry of Education. One study has been accepted and in one the intervention has been deferred.
- **S. Africa** – Scoping study awarded for IEs of preschool & school feeding programmes, but the IA has plans to do the study itself.
- **Philippines** – RFQ underway with Department for Education, for PWPG to identify up to 3 areas for IE.
- **Sri Lanka** – the RFQ is out, pending approval from the Ministry of Foreign Affairs
- **Uganda Kulika** – the RFQ is out
- **NEPAD** – the RFQ is out

34. To sum up, 3ie has not yet succeeded in generating the demand-led impact evaluations which were a central objective of Component 1 of the 3ie strategy. Progress is beginning to be made but there is some way to go before it can be certain that it has identified a workable approach to meeting that objective.

2.3.3 Thematic Window Grants

35. 3ie's first thematic grant window, the Social Protection Thematic Window, was launched in October 2011. SPTW is supported by DFID, and the two organisations worked together to develop it: forming an Expert Advisory Group, carrying out a survey of policy makers and commissioning a scoping study to map SP interventions worldwide and identify key knowledge gaps and research questions. The study identified three areas for funding under the SPTW:

1. The overall impact of social protection programmes and long term effects.
2. Political economy and institutional analysis.
3. Design issues.

36. This more targeted approach to commissioning impact evaluations represents an important shift from the more flexible Open Windows. How it has worked is considered in detail in Section 5.2. A DFID/3ie review panel selected 10 grants for funding in March 2012, and recommended that 3ie might seek alternative funding for five more. Grant agreements have been completed for nine out of the ten.

2.3.4 Systematic Reviews

37. 3ie awarded its first round of seven systematic review grants in early 2009. Five were due to be finished that year, and two in 2010. Five are now complete, but three grants were allowed substantial extensions, to mid/late 2011, and two of those remain outstanding. These two have been allowed a further extension to mid-2012. Although the grantees have been slow, there have also been delays in the review process, five months in one case. On one grant, the poor quality of the quantitative analysis meant that 3ie subcontracted that part of the work to a separate researcher. The result was further delay in merging the two analyses together.

38. Only two grants were awarded under SR2, one in late 2009 and the other in mid-2010. Neither has been finished. For this round, 3ie asked researchers to register their work with either the Campbell or the Cochrane Collaboration. Although this is expected to guarantee that the work meets the highest possible standards, the processes are slow. For one grant it took a year and a half for the research protocol to be approved.
39. In systematic reviewing 3ie has established an important partnership with DFID, and more recently with AusAid. In 2010, the 3ie team managed a pilot DFID call for 20 Systematic Reviews. This led on to grants from DFID and AusAid to coordinate and manage a joint call DFID-AusAid-3ie call for SR proposals. Of 32 grants awarded, 3ie is funding six from its own resources, in effect as its grant SR3 window. SR4 is also a partnership, this time with NORAD, SightSavers and Populations Services International. The round was launched in 2012, with 8 accepted from 40 applications. SR5 will be in partnership with CIDA, AusAID and MCC. For these partnership calls, each member of the consortium puts forward the research questions it wishes to see subjected to systematic review.
40. An additional four Systematic Reviews were conducted internally by 3ie staff.

2.3.5 PPGs and Others

41. In order to provide LMIC researchers and policymakers with the assistance of an international researcher and encourage them to bid for 3ie grants, in 2009 and 2010 3ie offered Project Preparation Grants of \$5,000. Of 37 applications, 14 received grants and seven subsequently applied through the 3ie open window. However only one received an OW grant.¹ This caused some disappointment among the applicants. With little demand and too low a success rate, it was decided not to continue the programme. It has been replaced by Policy Window Preparation Grants.
42. In another initiative to involve LMIC researchers, 3ie awarded five Water Supply and Sanitation grants of between \$10,000 and \$20,000. By March 2012, three were largely complete. Although this was mainly intended as a quick-win product, all five are expected to produce a final paper and one peer reviewed paper has been published.
43. In 2010, an RFP was published for Experience Sharing grants, to allow practitioners to share their experiences with implementing and using impact evaluation. The response was disappointing, but seven grants were awarded and two have already resulted in 3ie Working Papers.

2.3.6 Summary

44. With what was initially a very small team, 3ie has put together a very substantial portfolio of IE and SR grants. Table 5 summarises financial disbursements to date and Figure 15 shows a timeline for the overall programme. With 86 major grants awarded, \$35.2 million committed and \$19 million disbursed it is clear that substantial progress has been made. The first large batch of IE grants is due to be completed in 2012. If that target is met, 3ie will have a solid grant management achievement to its credit. Progress has perhaps been slowest for what should have been relatively short grants for systematic reviews. It is indicative of the problems with this kind of grant that one of the nine SRs was completed by 3ie's own staff, two others needed significant support from 3ie and for one SR, external consultants had to be hired to ensure the calculations were robust.
45. 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. These more strategic developments are discussed in more depth in Section 6.
46. The establishment of robust grant management procedures is itself an important achievement. This is reviewed further in Section 5.

¹ 3ie PPG Internal Evaluation Draft Report

Table 5: Financial Status of 3ie Grant Portfolio - March 2012

Grant Monitoring Report as on 30th March 2012								
Window	Budget (US\$)	Number of Studies	Funds Disbursed (US\$)					Balance
			Cumulative 2009	Cumulative 2010	Cumulative 2011	(Jan'12 - Mar'12) Q1	Cumulative to date	
OW1	4,231,524	17	901,949	576,627	946,761	493,569	2,918,905	1,312,619
OW2	15,111,035	30	-	5,926,194	2,490,634	2,051,462	10,468,289	4,642,746
OW3	9,011,201	21	-	-	4,108,141	142,157	4,250,297	4,760,903
SR1	349,347	7	90,379	42,427	123,856	38,955	295,617	53,731
SR2	178,544	2	-	14,985	11,796	107,126	133,907	44,637
SR3	431,444	6	-	-	-	117,955	117,955	313,489
PW1	659,886	3	-	-	77,383	15,000	92,383	567,503
PW2	39,486	2	-	-	15,000	4,743	19,743	19,743
PPG	70,000	14	25,000	35,000	5,000	-	65,000	5,000
Other	123,048	12	13,600	49,274	25,960	-	88,834	34,214
Total Grants	30,205,515		1,030,928	6,644,506	7,804,530	2,970,966	18,450,929	11,754,586
Reviewer & Meeting Exp.			39,766	108,841	13,700	36,050	198,357	-
Grand Total	30,205,515		1,070,694	6,753,346	7,818,230		18,649,286	11,754,586

2.4 Policy Advocacy and Communications

47. Advocacy and Communication (A&C) are a critical element in the 3ie Strategy. As the Strategy documents puts it, *“the real challenge facing 3ie is not to support the production of quality impact studies, but to support the production of quality studies which have an impact on policy.”* Strategy Component 3 – Building a Culture of Evidence-Based Decision Making – is directly focussed on ensuring policy impact. However, all components have important elements of A&C. Although 3ie’s advocacy and communications efforts started even before its formal establishment, its programme is summed up in its 2010 Learning Strategy for Policy Impact which has been fleshed out in subsequent workplans. The six objectives are summarised in Table 6.

Table 6: 3ie Policy Advocacy and Communications Objectives

A&C Objective	Detail	3ie Strategy Components
Enhance the policy influence of 3ie funded research and help 3ie grantees to better communicate results	Communications support to grantees including PIPs, policy influence clinics, online toolkit, studies on policy influence (including PIM project)	2 & 3
Promote the use and uptake of policy relevant impact evaluations	Knowledge sharing events (conferences), dissemination and media outreach (website, external communications), establishment of C2E index	1 & 4
Position 3ie as a trusted ‘knowledge broker’ on IEs	KM strategy and emerging community of practice (still under discussion), collaborating on publications e.g. SRs, WPs, JDE, newsletters	1 & 4
Initiate evidence-based campaigns	Preparation of campaigns drawing on emerging evidence from 3ie-funded studies; building alliances	2 & 3
Mobilise funding for IEs to increase the demand for evidence-based policy	Contributing to 3ie’s fundraising strategy: building strategic alliances and increasing membership base	5
Enhance information flow management and internal communication	Internal communication framework and tools: developing website, managing listservs, internal and corporate communications	5

Source: Learning Strategy for Policy Impact, 2011 & 2012 Workplans

2.4.1 Development of the PACO Function

48. The strategy summarised above has been developed over the four years since 3ie was established. First steps included setting up the 3ie website, in Q2 of 2008, the major Cairo conference in Q1 of 2009 and other demand generation activities. These were led by the ED. What is now the Policy Advocacy and Communications division (PACO) began to take shape with the appointment of the first A&C officer in February 2009. Since then the division has grown to include two more A&C officers and, more recently, a Chief Policy, Advocacy and Communications officer.

49. 3ie's expenditure on A&C (i.e. Strategy Components 3 and 4) has risen substantially from \$537,000 in 2010 to \$1.8 million budgeted for 2012. This increase from 4.7% of the 3ie total to 12.2% represents a significant change from the 3ie Strategy 2010 – 2013, which only allocated 2.8% to Components 3 and 4 together: only \$1.8 million for the whole four years.
50. The first 18 months was spent developing systems and approaches, culminating in early 2010 with the presentation to the Board of the paper 'Learning for policy impact strategy: 3ie's evidence-based advocacy and communications'. Key activities have included:
- Communication: maintaining the website, dealing with branding and house style, finding suitable listservs to publicise 3ie's calls for proposals
 - KM: setting up knowledge management systems
 - Grants: setting up protocols for communication with grantees, and publicity for grant calls
 - Publication: Annual report; technical Working Papers; IE and SR reports; the Enduring Questions publication series, which has been replaced by the Evidence Matters series; and newsletters
 - Developing a strategy for internal communications with the Board and with 3ie members
51. Experience with the Open Windows, and in particular the first Policy Window, has underlined the importance of 3ie's advocacy and communications effort. Although 3ie staff have led the effort, they have commissioned three supporting initiatives from ODI:
- Assistance with the further development of the Knowledge Management strategy
 - A series of Policy Impact clinics for policy-makers and implementing agencies
 - A large Policy Influence Monitoring (PIM) project
52. Table 7 overleaf summarises the history of 3ie's PACO effort since 2009.

Table 7: PACO History - summary to date

		Main PACO activities		Publications						
				WP	SR	EQ	EM	PB	NL	JDE
2008		ED set up the website, launched OW1 and the WP series, convened the Cairo conference	Focus on establishing the basics of communications & KM							
2009	Q1	A&C officer joined Supported advocacy work for OW1 Kuwait workshop to identify EQ themes WP series launched		1		5		5	1	1
	Q2	Support to launch of OW2 Delhi seminar series launched		2		3		1	1	1
	Q3	Support to launch of SR2 SR series launched		1	1			2	1	1
	Q4	Drafting of impact and learning strategy		1	1	2		4	1	1
2010	Q1	Communications Officer joined Impact & Learning Strategy presented to Board		2				1	1	1
	Q2	Support to launch of OW3		1				2	1	1
	Q3	LIDC London seminars launched			1			2	1	1
	Q4	Support to launch of SR3 ODI project commissioned MOU drafted with IDS							1	1
2011	Q1	On-going support to grantees		3					1	1
	Q2	Cuernavaca conference & policy clinic	1	1			1	1	1	
	Q3	KM strategy commissioned from ODI Support to launch of SPTW PIM project commissioned	1	1				1	1	
	Q4	Head of PACO joined On-going support to grantees	1					1	1	
2012	Q1	Support to launch of SR4 Comms support officers joined On-going support to grantees				1		1	1	
	Q2	Website revamp Support to launch of OW4 Rome policy influence clinic		1				1	1	
	Q3	Sri Lanka policy influence clinic								
Totals to date				15	6	10	1	18	14	14

Note: WP: Working Papers / SR: Systematic Reviews / EQ: Enduring Questions / EM: Evidence Matters / PB: Policy Briefs / NL: Newsletter / JDE: Journal of Development Effectiveness

2.4.2 The Policy Impact of 3ie Grants

53. With 80% of its funds expected to go to the production of evidence through IEs and SRs, 3ie can only achieve policy impact at scale if the studies funded are policy relevant and accessible to policy makers. Table 7 above summarises the way Advocacy and Communications are integrated into the grant process to try to make sure this is the case.
54. PACO’s role in this has become steadily more important. For OW1, PACO advised on the review process and on the guidelines for reviewing applications. Successful applicants were asked to provide a communications strategy though many grantees paid minimal attention to it and allocated no budget to dissemination. For OW2, a template for a project communications strategy was provided. However, this was too prescriptive and generally poorly done. This led to the concept of a Policy Impact Plan (PIP) which was tested with one OW1 project. This pilot was deemed successful and the PIP has become mandatory for OW3 and the SPTW, with a modified version being used for the SR series. OW1 and OW2 grantees were given the option of creating a retrospective PIP, which several did. There are currently 67 PIPs in total: 10 from OW1, 25 from OW2, 22 from OW3, and 10 from SPTW. They are reviewed in detail in Section 6.
55. Later in the process, 3ie also promotes the results of completed grants. On receipt of the draft final report, PACO works with the grantees to develop a communications strategy. This varies depending on the quality of the study and the grantees’ communications ability. On receipt of the full final report, PACO develops a separate communications strategy focused on what 3ie will communicate about the study. Where projects are innovative (e.g. REAP) they have been encouraged to develop their own communications products and asked to blog about their experiences; others have need a good deal more support from 3ie staff.

Table 8: Completed Projects - Communications and Advocacy outputs²

Project no	Short title	Final report	Workshops & presentations	Policy briefs / working papers	Newspaper articles	Misc. project briefs & flyers	JDE tech briefing	Journal articles *
OW1.66	REAP: Anemia in China	1		2				1S (BMJ)
OW1.69	Ghana cook stoves	1				1		
OW1.7	No Margin No Mission?	1				2		
OW1.78	Food supplementation in an AIDS programme, Uganda	1	3					1P (JDE)
OW1.9	The Promise of Preschool in Africa: ECD in Mozambique	1	6		3	3	1	

**S indicates submitted, P indicates published*

² Source: 3ie data on deliverables by tranche payment

56. 3ie has put an increasing amount of effort into promoting the policy impact of its grants. The 2010 Advocacy and Communications Strategy indicated a need to help grantees improve their communication skills to achieve greater policy impact. To do this, 3ie commissioned ODI to facilitate a series of Policy Influence Clinics. So far three have been held (Cuernavaca, Rome, and Colombo) and a fourth is planned for Africa. The two-day clinics bring together 3ie grantees (research teams, programme managers) and their policy counterparts, to provide training in communications techniques, to work on the PIPs (particularly indicators of impact), and to share preliminary results and challenges between the teams. Complementing the clinics is an online toolkit³, based on ODI's ROMA approach to improving policy influence. The clinics and toolkit are assessed as part of the output review (PIP process). The importance of distinguishing between policy relevance and policy impact is discussed in section 6.
57. The second initiative relating to policy impact is the Policy Influence Monitoring project (PIM).⁴ This is also managed by ODI working with three regional think tanks: CEPA in South Asia, CIPPEC in Latin America and CommsConsult in Africa. The project, which began in late 2011 and will last three years, will monitor and analyse the stakeholder engagement by 3ie grantees, to develop a detailed understanding of how to maximise the likelihood that impact evaluations will affect policy. All projects will have their PIPs reviewed. Between five and eight projects will be chosen for intermediate level monitoring and another 5-8 for more intensive monitoring. A project such as this will provide useful information on the pathways to impact for impact evaluations: there are no other empirical studies following research projects at this level of detail. The results will feed into both 3ie and the wider community interested in improving policy impact. However, care will need to be taken, not to micro manage grantees' communications and advocacy work, nor (more importantly) to blur the line between monitoring likely policy impact and advocating for particular policies – a point taken up in section 6.
58. In the 2012 work plan, \$0.148m is set aside for direct policy monitoring (of the PIPs) and \$0.407m for managing policy monitoring activities (the PIM project).

2.4.3 Publications

59. PACO leads the preparation of several different types of publication.
- Working Papers on conceptual and technical issues relating to impact evaluation; written by 3ie staff or by external authors. To date 13 WPs have been produced, each with a print run of between 500 and 1500.
 - Systematic Reviews are published upon completion: the five that have been published to date have had print runs of between 500 and 750.
 - Enduring Questions / Evidence Matters: The Enduring Questions (EQ) series was a response to the objective set out in the Founding Document for strategic programme to address such questions. Partly because of a Board decision not to pursue Thematic Windows (see the discussion elsewhere in this report), the series has been replaced by the

What can a policy brief achieve?

Using a randomised trial, IDS and 3ie investigated the 'stickiness' of messages from a systematic review under three circumstances: (a) without an accompanying opinion piece, (b) with an anonymous opinion piece, and (c) with an opinion piece from someone well known. Results showed that policy briefs do not change people's opinion, but they do shape beliefs for people who did not have a prior opinion. Accompanying the brief with an opinion piece by a sector expert increases the likelihood that people will share the information.

³ Currently hosted at <http://policyimpacttoolkit.squarespace.com/> It went live in June 2012.

⁴ Information taken from ODI/CIPPEC/CEPA/CommsConsult proposal document.

Evidence Matters policy brief series which distils evidence from systematic reviews. It is produced in partnership with IDS: one was published in Q2 2012, the other is in preparation. EQ briefs had print runs of between 250 and 500.

- Newsletters – a regular e-newsletter has been produced approximately quarterly from 2009 onwards, and sent out to the expanding 3ie mailing list. The newsletter currently reaches approximately 100 listservs, portals, associations and other newsletters). It is a mixture of information on 3ie and IE-related news from other organisations.
- 3ie briefs – these are short briefing papers on specific issues, based on work done by 3ie and other organisations. 14 briefs were published in 2009; a further brief has been published in 2012.
- The Journal of Development Effectiveness is a joint initiative by 3ie and Taylor Francis, published quarterly. 3ie’s ED is the Managing Editor, and the editorial board consists of nine leading impact evaluation specialists. The journal is a mix of original articles, reports from SRs, articles specifically on IE design, themed sections and themed issues. In general each issue contains between 5 and 9 articles and combinations of comment, reports of SRs and miscellany.

60. The 2012 work plan notes a budget for the different types of publication of \$0.152m, split between publications containing the results of specific impact studies (70%), policy briefs (17%) and conceptual papers (7%). The remainder of the budget will be spent on dissemination of the various publications.

2.4.4 3ie Events

61. 3ie has hosted two major conferences: the first, in Cairo in April 2009, was the first truly international conference on impact evaluation. It was co-hosted with AfrEA (which mobilised the evaluation community) and NONIE (which helped with fundraising). The second conference was held in Cuernavaca in June 2010. Data on feedback from the Cairo conference survey showed a good level of appreciation for the conference, though there were a few comments about the need to involve more developing country speakers.

Table 9: Cuernavaca Conference Expenditure (460 participants)

Item	Income US\$	Expenditure US\$
Donor support (AusAid, IDRC, IDB, World Bank, MCC)	383,830	
Conference fees	104,880	
Total receipts	488,710	
Local costs (to local partner)		337,439
Other costs		123,763
Total costs		461,202
Balance	27,208	

62. Nearly 60% of the conference expenditure was on airfares and bursaries for speakers and others attending the conference. In all, a quarter of participants were on a 3ie bursary: 103 out of 460.

63. The Cuernavaca conference objectives were set out as to:

- **Bridge the gap between researchers who are evaluating development intervention, and policy makers and program managers who are designing and implementing those programs.** There was a good representation of policymakers (308 out of the 405 who attended) suggesting that the mix was appropriate, though it is not clear whether or not this resulted in any lasting bridges being built between any of them.
- **Provide a platform to share and discuss experiences on how to best achieve evidence-based policy in sectors that are highly relevant for Latin America and other regions.** The conference sessions and 9 pre-conference clinics covered various methodological issues (the clinics allowing more detailed exploration).

64. A survey sent to the Cuernavaca participants 3 months after the conference (with a response rate of 39%) 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 programmes⁵.
65. It is not possible to estimate staff time spent on organising, running and evaluating the conferences; that breakdown only exists for 2011 onwards. While all conferences are a significant drain on staff time, the two 3ie conferences do appear to have served the purpose of establishing (Cairo) and maintaining (Cuernavaca) 3ie’s profile in the international evaluation community. There was significant online traffic during the conference: its promotional material was cross-posted on 40 websites of other organisations, there were 4,400 unique visitors to the website in June, and 32 blogs which generated 3,584 visits).
66. 3ie supports seminar series in Delhi, London (with LIDC) and Washington (with IFPRI). These are to 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. The target is one seminar per month, though it depends on whether interesting people can be found to speak. Topics are therefore decided in a relatively ad-hoc manner⁶. It is not possible to develop a strategy or themed approach, as it is too dependent on who is available. The table shows the number of seminars in each centre: the total cost of the seminars to 3ie is estimated in the 2012 Workplan to be \$0.023m.

Table 10: 3ie Seminar Series

Location	No of seminars			
	2009	2010	2011	2012 (to date)
Delhi	7	7	6	2
London	0	0	9	7
Washington	0	0	0	2

2.4.5 Website and corporate communications

67. The website has always been seen as a central part of the A&C strategy; not only as the outward face of 3ie but also as a major communications tool. Via the website, PACO publicised calls for proposals, informs the IE community of funding opportunities / workshops / useful publications, disseminates the newsletter, and acts as the repository of expert information such as the expert

⁵ Source: Cuernavaca post-conference survey report, on website.

⁶ From interviews with the London office

roster, IE database, useful resources, and of 3ie-generated studies, working papers and other publications.

68. The communications team draws heavily on 3ie's database, using it to write features analysing the findings from SRs and IEs. They also invite grantees or other experts to write blogs, and are expanding into social media. Much dissemination is electronic; the team continues to maintain a long list of IE-related listservs and uses them to target dissemination activities. The new website was launched in June 2012 following a user analysis; it has been redesigned to be more user-friendly to policymakers who were discovered in interviews to approach issues sectorally. The cost of the website design was £26,500 (\$41,000). This appears to be separate from the total budget for website management of \$0.042m as set out in the 2012 Workplan.

2.4.6 C2E (Commitment to Evaluation)

69. As part of component 3 of the 3ie strategy (developing a culture of evidence), 3ie launched the Commitment to Evaluation (C2E index) in November 2011. Following a first consultation in Cuernavaca, a survey of donors, a rapid poll in Busan and internal discussions, a paper has been commissioned to set the scope of a possible initiative and identify possible partners. A working group has been established with donor and policy maker representatives and their report is due shortly. The emerging idea is to move away from a specific evaluation index and instead to pilot two initiatives. The first would set up a peer learning mechanism (similar to that used by OECD DAC) for 3ie's members to share experiences about getting evidence into use and for more formal discussions to be held at members' conferences. This may also go some way to addressing the issue of the benefits members gain from membership of 3ie (see section 2.5.3), though cost will be an issue. The second is collaboration with UNDP to develop an indicator of the use of evidence which they would implement with their country programmes. \$0.018m has been budgeted in 2012 for C2E activities.

2.4.7 Conclusion: 3ie's Advocacy and Communication Achievements

70. 3ie's advocacy and communications effort has developed substantially since 2009. A considerable amount was achieved with a very small team; this has recently expanded and allowed a greater focus on specific communications issues. On the communications side, a wide range of publications is produced; some at regular intervals (newsletters, Journal of Development Effectiveness articles, newsletters) and others as results become available (Working Papers, Systematic Reviews, the EQ/EM series). The website was substantially revamped in 2012 following a consultation on what evidence stakeholders were looking for and how they searched for it. In terms of policy advocacy work there has been a similarly large effort to design and implement the PIP process and support grantees via clinics and the online toolkit. The development of both strands of work was informed by comprehensive strategy development processes.
71. There remain several questions about how to ensure that 3ie grants are policy relevant and how to build a culture of evidence-based policy making. These are discussed in more detail elsewhere. There have also been changes in the A&C strategy which raise strategic issues, such as the shift from Enduring Questions to Evidence Matters and the decision to downplay attempts to build a community of practice. These will also be considered further elsewhere.
72. PACO has developed a clear strategic approach to strengthening the policy impact of 3ie's work and important initiatives launched to implement it. However it is not immediately clear why the budget for PACO activities has grown so significantly from that set out in the 3ie strategy 2010-2013. As set out in section 6, there remains a question of whether this focus on policy impact should remain a pillar of 3ie's overall strategy.
73. That said, it is clear that in PACO 3ie has the effective advocacy and communications unit it needs to support whatever strategy the executive and the Board decide. Table 11 presents a summary

of the A&C products and events that 3ie has produced since its establishment. It is a snapshot of a substantial achievement.

Table 11: Summary of 3ie Publications and Advocacy Efforts

	2008	2009	2010	2011	2012*
Events					
Conferences	1			1	
Seminars		7	7	15	12
External Publications					
Enduring Questions / Evidence Matters		10	0	0	1
Working papers		5	3	6	1
Systematic Reviews		2	1	2	1
Newsletters		4	4	4	4
PIPs		17 (OW1)	30 (OW2)	22 (OW3)	10 (SPTW)
Policy Briefs		10	5	1	1
Journal of Dev. Effectiveness (vols (issues))		1(4)	1(4)	1(4)	1(2)
Internal publications					
Annual reports		1		1	
PACO strategy documents			1	1	
Downloads					
Visits to website (avg daily)**		167	175	238	257
Downloads (avg monthly)¥					
• WPs				478	681
• SRs				68	726
• PBs				40	335
JDE downloads					

* To date

** Detailed web analytics only available from mid-2011 on. Other figures imputed from ED reports

¥ Average monthly figures, imputed from web analytics data. Details not available pre June 2011

2.5 Feedback from 3ie Applicants and Others

74. 3ie has shown itself to be a reflective organisation. It has made a good effort to get feedback from stakeholders at appropriate points. This includes a survey of Associate Members, surveys taken after the Cairo and Cuernavaca conferences, and surveys of applicants for grants under OW2 and OW3. Through ODI a further survey of 3ie grantees was conducted, to gather views to inform their KM strategy. There has also been a NONIE survey on the role of 3ie. Table 12 summarises the response rates and scale of these surveys. The rest of this section briefly summarises the results of those available to the Evaluation Team. For comparison, the table also shows the same data for the three surveys carried out by Triple Line for this Process Evaluation. The results of these are presented separately, in Chapter 4.

Table 12: Stakeholder Surveys carried out by 3ie, Triple Line and others

SURVEY	DATE	Response Ratio	No. of Questions	Report published
1. OW2 applicant Survey	Jan-10	62/255	9	yes*
2. OW3 applicant Survey	Oct-10	121/366	25	yes*
3. Cairo Conference	Mar-09	134	12	no
4. Cuernavaca Conference	Sep-11	139/460	19	yes*
5. Associate members survey	Apr-12	23/93	8	no
6. ODI grantee Survey on KM strategy	Mar-11	35/68	65	no
7. NONIE survey of 3ie	Mar-11	28	69	no
8. Triple Line - grantee PIs	Jun-12	36/88	51	Section 4.2
9. Triple Line - grantee CNs	Jun-12	21/45	30	Section 4.1
10. Triple Line - 3ie Stakeholders	Jun-12	409	40	Section 4.3

Note: * Published on 3ie Website

2.5.1 Applicant surveys

75. These online surveys had questions on the satisfaction of applicants, how they had heard of 3ie and the usefulness of the assistance 3ie had given to applicants. Summaries of the results are published on the 3ie website, as a one page summary of what applicants liked and didn't like, and a further page on what 3ie will do with the findings. Overall applicants were satisfied, and rated 3ie's grant process as the same or better than other grant makers. Shortcomings came from response times of queries, and a perception for some that the process is not always transparent or fair.
76. The OW3 applicant survey was somewhat longer than for OW2, with 23 multiple choice questions, and eight textboxes. The response rate was slightly better in the second survey - 33% versus 24%, and 75% of the respondents had not received an award. Satisfaction rates were high, but somewhat lower for OW3. For its third open window, 3ie introduced an online Grant Management System (GMS) and a two-stage application process. 95% rated the GMS as good (47%) or satisfactory (48%), although there were a number of comments on how it could be improved. 88% felt the two-stage process was a good idea. Response times from queries seemed to have improved upon since OW2, but transparency in reviewer feedback, and feedback in general, especially in terms of reasons for rejection still seemed to be an issue.
77. Overall the applicant surveys drew out useful feedback and the results are fairly presented in 3ie's reports, which note the potential biases and weaknesses. One criticism could be in that the OW3 report does not make any comparison with OW2, to see how things had changed. All in all the applicant surveys seem to have succeeded in creating a dialogue between applicants and 3ie management, which has drafted action points to respond to the criticisms. These have been acted upon, with improved reviewer feedback reports, and an introduction of a query response system in between the grant application deadline and the award announcement in response to action points from the feedback of the surveys for OW3 and OW2.

2.5.2 Conference Evaluations

78. Over 700 people attended the 2009 AfrEA-NONIE-3ie Impact Evaluation Conference in Cairo. The post conference survey asked about the usefulness of different aspects of the conference. The findings of this survey were not reported, but the evaluation team has reviewed the data. This shows large majorities (80-90% 'strongly agree' or 'agree') found the conference relevant and

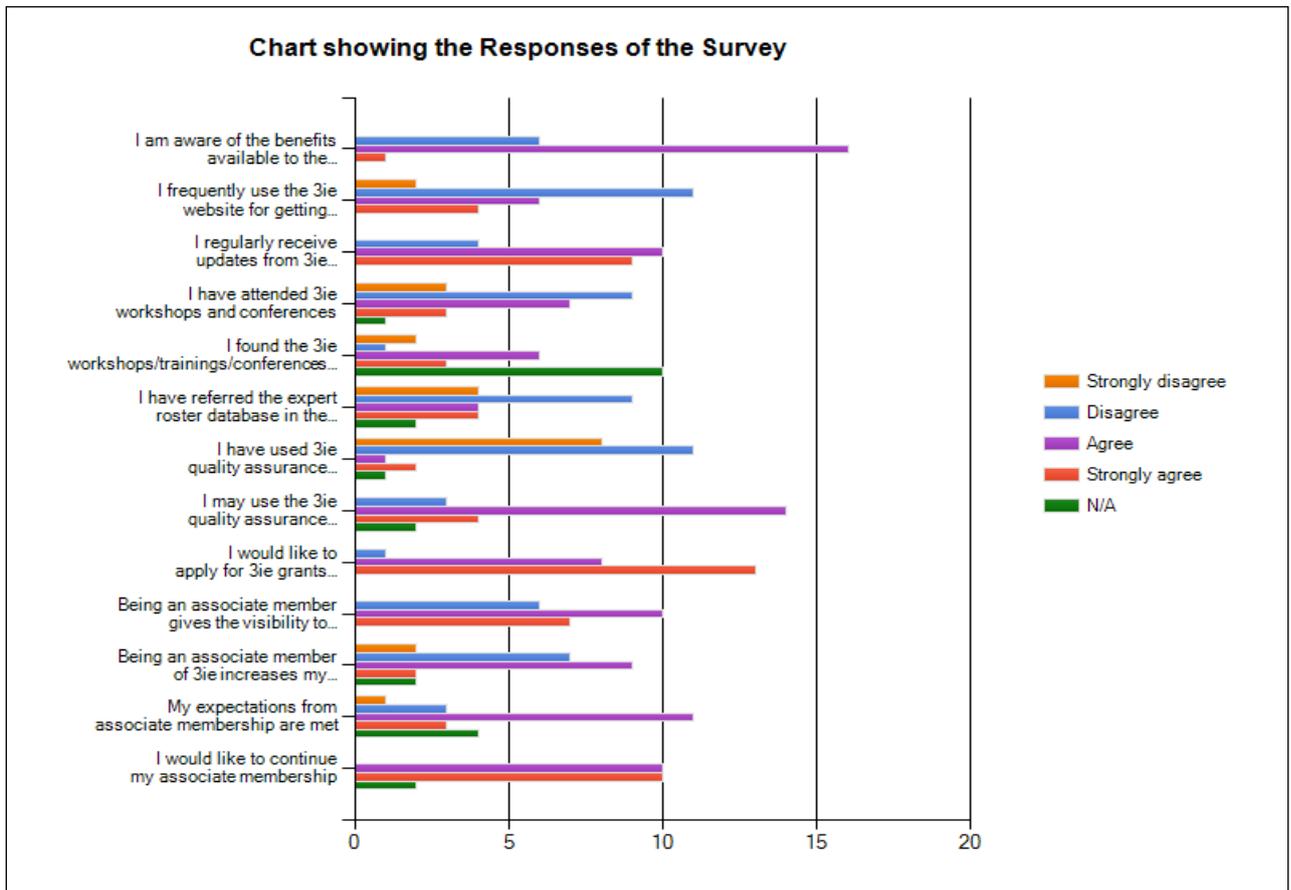
likely to inform their impact evaluation work. Without a published report, it is hard to know whether lessons were learned, and comments used to inform future conferences.

79. A similar survey was conducted of the 460 participants to the 2011 3ie Mind the Gap Conference in Cuernavaca. As for the OW2 and 3 surveys, summary findings - What 3ie does well, what 3ie needs to improve upon etc. – are published online. The survey had more, and more specific, questions than the Cairo survey. Of the 139 respondents many chose not to answer several questions. Nevertheless, the feedback has been well presented, with criticisms addressed set out and published, to be acted upon for the next conference.

2.5.3 The Associate members survey

80. In 2011 3ie carried out a survey of Associate Members' perception of their membership, to identify ways to constructively engage with them. The results were briefly presented in ED's report to the seventh meeting of the 3ie Board of commissioners.
81. The Evaluation Team have reviewed the data. Of 93 Associate Members, 22 responded, the majority from Research organisations (12) and academic institutions. A quarter of respondents were not aware of the benefits available to associate members, and more than half did not frequently use the website.
82. About half had attended workshops and events, and found them useful and relevant. The fact that 21 of the 22 agreed that they would like to apply for 3ie grants in the future perhaps reveals the AM's principle motivation for registering. Most also thought that being an associate member gives visibility to their organisation through listing on the website. Only a third had not used the expert roster, even fewer 3ie's quality assurance services; although most said that they may use them in the future.
83. Overall 14/22 thought that their expectations from AM were met, 3 disagreed and one strongly disagreed, with 4 choosing not to answer. All wanted to continue their associate membership. Figure 2 shows the chart of Survey Responses, from the 3ie report.
84. When asked what services 3ie might provide to associate members, the response was more opportunities to attend 3ie workshops and conferences, and shared survey instruments and access to datasets, and listings of other (non-3ie) grant opportunities. Associate members were keen to engage with other associate members, to collaborate through an active internet network, meetings and conferences to share research publications etc. The idea of creating communities of practice through associate membership was welcomed by some, though they expressed a lack of time and resources to do so. Other suggestions included sponsoring of associate members from developing countries to attend workshops and seminars and inviting associate members to at least a fraction of board and members meetings.

Figure 2: 3ie Associate Member Responses



85. With the exception of the Associate Member survey, response rates were reasonable if not high for 3ie’s surveys, and lessons learned have been written up and largely acted upon. No surveys were done of the Policy Window applicants or SR window applicants, however a survey of SPTW applicants is currently being undertaken.

2.6 Organisation and Staffing

86. Figure 3 presents the current 3ie organogram, showing its five operational divisions: three in Delhi, one in London and one in Washington.

Figure 3: 3ie Organogram



2.7 Structure and Staffing

87. 3ie’s distributed structure is built around hosting arrangements with GDN in Delhi and the London School of Hygiene and Tropical Medicine. Staff in Delhi are recruited on GDN contracts, on GDN terms and conditions. Staff in London are contract through LSHTM. Only the staff in Washington are contracted directly by 3ie.
88. After it was incorporated in 2008, 3ie had no staff until the end of that year. Howard White had been appointed Executive Director, but he was contracted through CGD as was William Savedoff who acted as Secretary-Treasurer part-time. By the end of the year, one Evaluation Officer had also been appointed in Delhi and two consultants were employed to work on SRs in London, on contract (subsequently becoming evaluation officers). The ED transferred onto a 3ie contract on 1 January 2009, and he was joined by his Deputy, a Senior A&C Officer, a Programme Officer and two assistants. The evaluation officers based in Washington and London continued on contract.

89. The 3ie team continued to grow, with seven more posts in 2010 and a further eight in 2011. In 2012, two more members of staff have joined and four more posts are to be filled. As discussed below, financial processes were largely outsourced to GDN, with 3ie's own accounts managed by a US accounting firm. 3ie did not appoint a Finance Officer until October 2010 and the Deputy Director Finance and Administration only joined 3ie in September 2011.
90. In some ways, 3ie has grown from the bottom up. By the end of 2010, there was a relatively strong group of Evaluation and Communication Officers supported by Research Assistants and Programme staff. If not reporting directly to the ED or his deputy, most of the team were in direct contact with senior management. The more recent appointment, between September 2011 and June 2012, of four Deputy Directors is likely to change the balance of the team.

2.7.1 Hosting Arrangements - GDN

91. Two documents cover 3ie's hosting by GDN: An Agreement setting out the overall principles and a Memorandum of Understanding on more detailed arrangements. The Agreement and the original MoU were signed in November 2008. A revised version of the MoU was signed in September 2010, incorporating two earlier addendums, but still based on the original agreement.
92. The Agreement envisages a close partnership between the two organisations:

"The Parties shall cooperate and collaborate closely in pursuing their shared objectives"

"The 3ie activities and programs shall be fully integrated into and form part of the organizational framework of GDN"

GDN's Board of Directors are to "maintain general policy oversight" over 3ie programmes and the President of GDN is to "exercise broad oversight over all 3ie activities, programs and staff."

93. The 2008 MoU implemented these principles by creating a GDN International Impact Evaluation Department (IIED) of six 3ie seconded staff. The ED was to hold the GDN rank of Senior Director. Four more GDN staff in programme management, accounting and administrative functions were dedicated to support the IIED and the 3ie programme.
94. The financial arrangements set out in the MoU sum up the arrangement. There are four headings for 3ie payments to GDN for its support. Table 13 sums these up for the first and latest MoUs.
95. For its annual fee, plus the Expansion payment, 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.

Table 13: 3ie Payments to GDN

	2008 MoU	2010 MoU
1. 'Pass Through'	- Salaries/staff costs - Travel - Event organisation - 3ie Grants	Unchanged
2. One-off Expansion charge	\$7,500 per new staff position	\$8,000 per new position
3. Annual Fee	Fixed Fee: \$368,000 PLUS on grants disbursed: > \$3 million pa – 8.0% > \$6 million pa – 5.6%	Fixed Fee: \$600,000 PLUS on grants disbursed: > \$4 million pa – 7.0% > \$6 million pa – 6.0%
4. Reimbursable	Any other agreed costs	

96. Apart from some adjustments to the scheduling of payments, the 2010 MoU includes two important changes:

- The number of 3ie staff was increased from six to 17 immediately, with provision for further increases to a maximum of 22.
- The inclusion of clauses to indemnify GDN against the risk that 3ie would not be able to fund its grant commitments. Given that it is the contracting party on the grant agreement, this is a significant risk for GDN.

97. 3ie's Founding Document sets out a vision of a membership organisation with strong southern voice. As part of that, there was early interest in the possibility of winning the status of an International Organisation (IO) for 3ie; a route which GDN had already followed. This, together with its established systems for awarding grants to research institutions, in particular southern research institutions, made GDN seem a natural partner for 3ie. Nevertheless, the Evaluation Team have been informed that the early years of the relationship were difficult. It had been expected that GDN would host 3ie in Cairo. However, it turned out that GDN did not have IO status in Cairo and so could not host 3ie effectively from there. There were also difficulties over the financial terms and a feeling, on 3ie's side at least, that GDN was not providing a good service. Long delays in the award of grant agreements for OW1, effectively the launch date for 3ie, were partly the result of difficulties with GDN's contract formats.

98. The relationship is now reported to be working better. One factor in that improvement seems to have been the development of 3ie's own team, in particular the establishment of its own grant management and financial control functions, and the definition of a firm division of labour 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 expectation in the original agreement that *"3ie activities and programs shall be fully integrated into and form part of the organizational framework of GDN ..."* The implications of this, and the future of the GDN hosting arrangement are considered further in Section 5.6.3, along with an assessment of how far the arrangement represents value for 3ie's money.

2.7.2 Hosting Arrangements - LSHTM

99. 3ie's hosting arrangement with the London School of Hygiene and Tropical Medicine is simpler. The first three-year agreement, signed in October 2010, was to employ two evaluation officers and three research assistants on 3ie's behalf. Apart from salaries, and a 10% management fee on the salaries, 3ie only has to reimburse the cost of small items LSHTM buys for it. The salary element

carries an annual inflation factor of around 6%. In November 2011, the agreement was amended to add two more staff posts, including a 3ie Deputy Director.

100. At £45,000 for a Research Assistant, some of the salaries charged under the agreement seem high. £26,000 to £30,000 would be normal at that level. However, this 'salary' figure represents the full cost of employment, including allowance for office space; the positions are on the LHSTM pay scales as rated by their job descriptions. Greater clarity on the composition of gross salaries would be useful.
101. Before the addition of two staff, the total charge in the first year of the agreement was £274,000, approximately \$430,000. From 2012, this will rise to over \$600,000 a year. 3ie staff are housed in the London International Development Centre, with an annual office cost of approximately \$40,000.⁷

2.7.3 The Washington Office

102. The evaluation team do not have details of the Washington office – leases, staff contracts etc. It is understood that a small office has been rented for the five staff (FTE) in post. The salary/benefit cost from 2012 onwards will be over \$700,000 a year, with annual office costs projected at \$20,000.⁸

2.8 Management Systems

103. 3ie's grant management processes are discussed in detail in Section 5.2 below. This section concentrates on planning and budgeting system and on HR management.

2.8.1 Planning, Budgeting and Financial Control

104. Until late 2010, 3ie had no dedicated finance staff. It was nearly a year later when the one Finance Officer was joined by the Deputy Director Finance in September 2011. These appointments were paralleled by the establishment in mid-April of a stronger Audit and Finance Committee (A&FC) on the 3ie Board of Commissioners and the appointment, in mid-2011, of a financial specialist as Chair of the Committee. In April 2012, the A&FC approved a 3ie Accounting Policies and Procedures Manual.
105. Prior to these developments, the evaluation team understand that the only formal statement of financial policies is that set out in the two and a half page Addendum to the 2008 Founding Document, which includes a brief statement on the delegation of financial authority to the Executive Director. In this earlier period, financial aspects of 3ie grants were managed by GDN. Other costs were paid from the US bank account, with the ED and Deputy ED as signatories for larger amounts and the Washington-based Chief Evaluation Officer for smaller sums. While there has always been reference to the table of approval thresholds, with senior management travelling and the ED only in Delhi on average every two months, this meant a significant amount of business was dealt with by email. 3ie's accounts were maintained remotely by a US accounting firm.
106. While there is no suggestion of any financial losses, it is clear that during its first two years of operation, 3ie was some way short of a fully integrated planning, budgeting and reporting system. The Board did not approve the Workplan and Budget for 2010 until October in that year. The minutes record that they asked for the 2011 Workplan to be more detailed: broken down by sub-activity, and with activities 'clearly linked to budget and staff time requirements'.⁹ On the

⁷ Data from 3ie Financial Projections, Board Pack – 7th Board Meeting

⁸ Ibid

⁹ 3ie Management report that the Board were presented with a detailed breakdown 2010 and that the Board's requirements were difficult to interpret.

accounting side, the audited Financial Statement for 2009 was not received until December 2010. Although the 2010 FS was completed on time, an auditors' Management Letter identified one material weakness, over the segregation of financial duties, and a number of significant deficiencies.

107. With a structured 2011 Workplan and timely Financial Statements for 2010 and 2011, the Board of Commissioners was able, for the first time, to review detailed reports on 3ie's financial performance against budget at its 7th meeting in March 2012.¹⁰ The Board pack also included an auditors' 2011 letter confirming that the prior year's material weakness and most of the significant deficiencies had been satisfactorily addressed.
108. At the time of the evaluation team's visit to Delhi, in May 2012, the finance team had transferred 3ie's accounts from the US contractor to an internal system, and they had just begun to issue Monthly Financial Reports. However, these are only simple statements of 3ie's financial position on a cash basis. This is a long way short of a full set of management accounts, by activity or cost centre.
109. Measures are in place to move towards more robust budgeting and financial reporting. A chart of accounts has been available since 2009. Staff record their time against the different activities. The new Accounting Manual includes a requirement that 'Budget Managers are assigned to each budget as it is approved and held accountable for expenses incurred'. This will only come fully into force from 2013, when it is understood that the Deputy Directors will be expected to draft budget proposals for their divisions and be 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. Once this system is operational, with monthly or quarterly reports against plan and budget, 3ie will have a complete planning, budgeting and financial control system in place.

2.8.2 Human Resources

110. 3ie staff are contracted under five different arrangements:
 1. International staff contracted through GDN on international terms (4, including ED)
 2. Indian staff contracted through GDN (13)
 3. GDN Delhi staff seconded to 3ie by GDN (3)
 4. UK-based staff contracted through LSHTM (7)
 5. US-based staff contracted direct by 3ie Incorporated (5)
111. In principle, the GDN Delhi Manual of Staff Policies and Service Rules covers all staff based in Delhi, including those on expatriate terms. The Manual sets out a comprehensive set of procedures: performance management, discipline, code of conduct etc. In practice, however, it appears that some aspects are not being applied to staff contracted on behalf of 3ie. Bonus arrangements and annual performance reviews are two significant examples, with 3ie carrying out its own performance reviews within the GDN timetable and applying their performance related pay increases and bonuses. The ED is the exception to this. London staff are governed by the policies and regulations of LSHTM and those in Washington by the terms of their individual contracts. So far 3ie has not developed procedures to define how it works within each set of conditions, or to bridge across the different arrangements. 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.

¹⁰ 3ie Management record that the Board had been given more detail in earlier years but it is not apparent from the minutes.

112. For specialist positions, especially those in the Evaluation and Systematic Review Offices, 3ie is concerned to attract staff of the highest international calibre. Even on international terms, 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. It is understood that there is pressure from the Government of India to keep the number of international posts to a minimum. As is discussed below, this may become an even bigger issue if it is decided to establish an independent 3ie office in Delhi, without GDN's status as an International Organisation.
113. 3ie is keen to encourage technical staff to further their careers. Key technical staff members are to be allowed 20% of their time for personal research. Staff based in London fall under the HR policies of LSHTM and London University and the majority of staff have academic status. This brings a requirement that they meet minimum teaching and publishing obligations. While obviously an important attraction for capable researchers, these independent academic activities may need careful management to prevent them weakening the focus on 3ie's own tasks.
114. 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, albeit remotely for much of the time.
115. As has already been mentioned, the overdue replacement of the Deputy Executive Director and the recent appointment of four other departmental heads will change this balance. The wider implications of the staffing structure are discussed below in Section 5.6-5.7.

2.9 Funding Status

116. Table 14 presents a summary of 3ie's funding status, showing actual receipts over the period 2009 – 2012 and future funding committed in written agreements for the period 2013 to 2016.¹¹ Three points stand out:
- At present only one donor, DFID, has made a commitment to core funding beyond 2013,¹² 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.¹³
 - Only one NGO has maintained continuous membership: SCF.
117. Apart from the financial implications of the situation from 2012 onwards, the auditors have pointed out that 3ie must meet a Public Support Test to maintain its status as a Public Charity. The auditors have confirmed that the full amount of foreign contributions can be included in the public support calculation, meaning that it is unlikely that dependence on a sole, non-US donor would affect the test.

¹¹ The 2012 figure combines actual receipts and projected further funding for the remainder of the year.

¹² The grant agreement for the BMGF –HIV/AIDS project makes it clear that the funds are only available for that project, though these include costs of staff time, consultants hired for external review and associated costs and overhead. (The Evaluation Team has not seen the budget.)

¹³ It is understood that the CGD funding came originally from the same group of core donors.

Table 14: Funding Status of 3ie - 2008 to 2016

Name of Donor	Actual					Projections				Total		% Share
	FY 2008	FY 2009	FY 2010	FY 2011	FY 2012	FY 2013	FY 2014	FY 2015	FY 2016			
FOUNDATIONS												
Bill and Melinda Gates Foundation	4,000,000	-	5,000,000	5,000,000	4,000,000	3,000,000				21,000,000		25.7
William & Flora Hewlett Foundation	-	-	2,000,000	2,000,000	2,000,000					6,000,000		7.4
Center for Global Development	5,150,000	350,303	-	-	-	-	-	-	-	5,500,303		6.7
											32,500,303	39.8
BILATERAL DONORS												
NORAD	100,000	100,000	99,970	-	200,000					499,970		0.6
Netherland Min. for Foreign Affairs	140,800		524,440	-						665,240		0.8
DANIDA	-	172,197		462,300						634,497		0.8
DFID	-	492,870	3,007,100	13,882,618	7,577,441	7,205,112	7,120,127	793,543	238,101	40,316,911		49.4
SIDA	-	136,122	-	154,892	264,855	301,071				856,940		1.1
AusAid	-	-	1,419,300	869,928	403,014	402,120				3,094,362		3.8
Millenium Challenge Corporation	-	-	100,000	-	100,000					200,000		0.2
CIDA	-	-	576,979	1,244,168	200,000					2,021,147		2.5
Irish Aid	-	-	97,733	-						97,733		0.1
USAID	-	-	-	100,000	162,500	225,000	62,500			550,000		0.7
											48,936,800	60.0
NGOS												
Save the Children	20,000	10,000	10,000	10,000	10,000					60,000		0.1
International Rescue Committee	-	10,000	-	-						10,000		0.0
Women for Women	-	20,000	-	-						20,000		0.0
Population Services International	-	-	-	-	30,000					30,000		0.0
Sightsavers International	-	-	-	-	30,000					30,000		0.0
Master Card Foundation	-	-	-	1,185						1,185		0.0
Higher Education for Development	-	-	-	-	8,333					8,333		0.0
Others	-	13,725	-							13,725		0.0
											173,243	0.2
Sub Total Core Funding	9,410,800	1,305,217	12,835,522	23,725,091	14,986,143	11,133,303	7,182,627	793,543	238,101	81,610,346		100.0
PROJECT FUNDS												
BMGF - HIV/AIDS	-	-	-	-	5,932,271	5,794,934	5,793,114	3,649,861	572,422	21,742,602		
Total Funding	9,410,800	1,305,217	12,835,522	23,725,091	20,918,414	16,928,237	12,975,741	4,443,404	810,523	103,352,948		

Table 15: 3ie Cashflows - Actual and Projected 2008 to 2016

	Actual					Projections				Total	
	FY 2008	FY 2009	FY 2010	FY 2011	FY 2012	FY 2013	FY 2014	FY 2015	FY 2016		
Carried Fwd from Previous Yr		9,299,978	7,142,522	8,605,678	22,420,759	23,106,901	25,232,960	23,408,343	15,194,642		
Core Funding Income	9,410,800	1,305,217	12,835,522	23,725,091	14,986,143	11,133,303	7,182,627	793,543	238,101	81,610,346	
Conference Revenues		882,839	101,006	776,683						1,760,528	
Total Funds Available	9,410,800	11,488,034	20,079,050	33,107,452	37,406,901	34,240,204	32,415,587	24,201,886	15,432,743		
Operating Expenses											%
- Salaries, benefits		373,431	525,529	1,506,536	2,000,000	2,000,000	2,000,000	2,000,000	2,000,000	12,405,496	16.1
- Professional fees		293,695	754,269	843,970	850,000	850,000	850,000	850,000	850,000	6,141,934	8.0
- GDN management fee		368,000	498,600	868,619	950,000	641,000	641,000	641,000	641,000	5,249,219	6.8
- Travel		325,109	333,954	437,072	500,000	500,000	500,000	500,000	500,000	3,596,135	4.7
- Conference costs		944,883	80,933	594,327						1,620,143	2.1
- Operational cost		16,353	105,394	255,525	500,000	500,000	500,000	500,000	500,000	2,877,272	3.7
Subtotal	110,822	2,321,471	2,298,679	4,506,049	4,800,000	4,491,000	4,491,000	4,491,000	4,491,000	32,001,021	41.6
Grants		2,024,041	9,174,693	6,180,644	9,500,000	4,516,244	4,516,244	4,516,244	4,516,244	44,944,353	58.4
Total Expenditure	110,822	4,345,512	11,473,372	10,686,693	14,300,000	9,007,244	9,007,244	9,007,244	9,007,244	76,945,374	100.0
Net Cash Flow	9,299,978	-3,040,295	1,362,150	13,038,398	686,143	2,126,059	-1,824,617	-8,213,701	-8,769,143		
Net Funds Carried Forward	9,299,978	7,142,522	8,605,678	22,420,759	23,106,901	25,232,960	23,408,343	15,194,642	6,425,499		

- Notes:
1. Core funding income from Table 14 above
 2. Expenditures 2008 – 2011 from audited Financial Statements;
 3. FY 2012 Operating Expenses and Grants rounded figures from 3ie Budget 2012, excluding contingencies
 4. FY 2013 – 2016 Operating Expenses as for FY 2012
 5. FY 2013 to 2016 Grants = Outstanding commitments (including OW4 and PW3) spread over 4 years
 6. It is assumed that the BMGF HIV/AIDS grant and any future conferences will have no net impact on 3ie's core funding balance

118. To provide an understanding of 3ie's funding situation, Table 15 presents a cash flow analysis from 2008 onwards, together with projections forward to 2016. Expenditure projections are based on 3ie's 2012 budget, commitments under current grant programmes and the proposed funding for OW4 and PW3 approved at the 7th meeting of the Board of Commissioners. This analysis 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. (The make-up of 3ie operating expenses is reviewed in Chapter 5.)

3 Governance, Direction and Accountability

119. After a brief statement of setting out the framework of 3ie’s constitution, this chapter seeks to assess how that framework is operating in terms of strategic direction, delegation of responsibilities, and accountability.

3.1 Constitution and Bylaws

120. 3ie has no formal constitution, but three governing documents serve the same purpose: the Certificate of Incorporation, the Founding Document, and the Bylaws.

3.1.1 The Certificate of Incorporation

121. 3ie was registered, on 14 May 2008, as a non-profit corporation in the US state of Delaware, to be “operated exclusively for charitable, scientific and educational purposes”, under section 501(c)(3) of United States tax law. This is the standard status for US NGOs and philanthropic foundations. It prohibits political activity, and provides that assets must be distributed to other non-profit organisations on dissolution. The Certificate of Incorporation names one specific purpose for 3ie: “(a) contributing to the fulfillment of wellbeing by encouraging and funding the research, production and use of evidence from rigorous impact evaluations for policy decisions that improve social and economic development programmes in low- and middle-income countries;” ... The only other purpose is a standard catch-all for doing “everything necessary, proper, advisable or convenient” for 3ie’s purposes and objectives. Under the terms of the Certificate, the Bylaws 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.

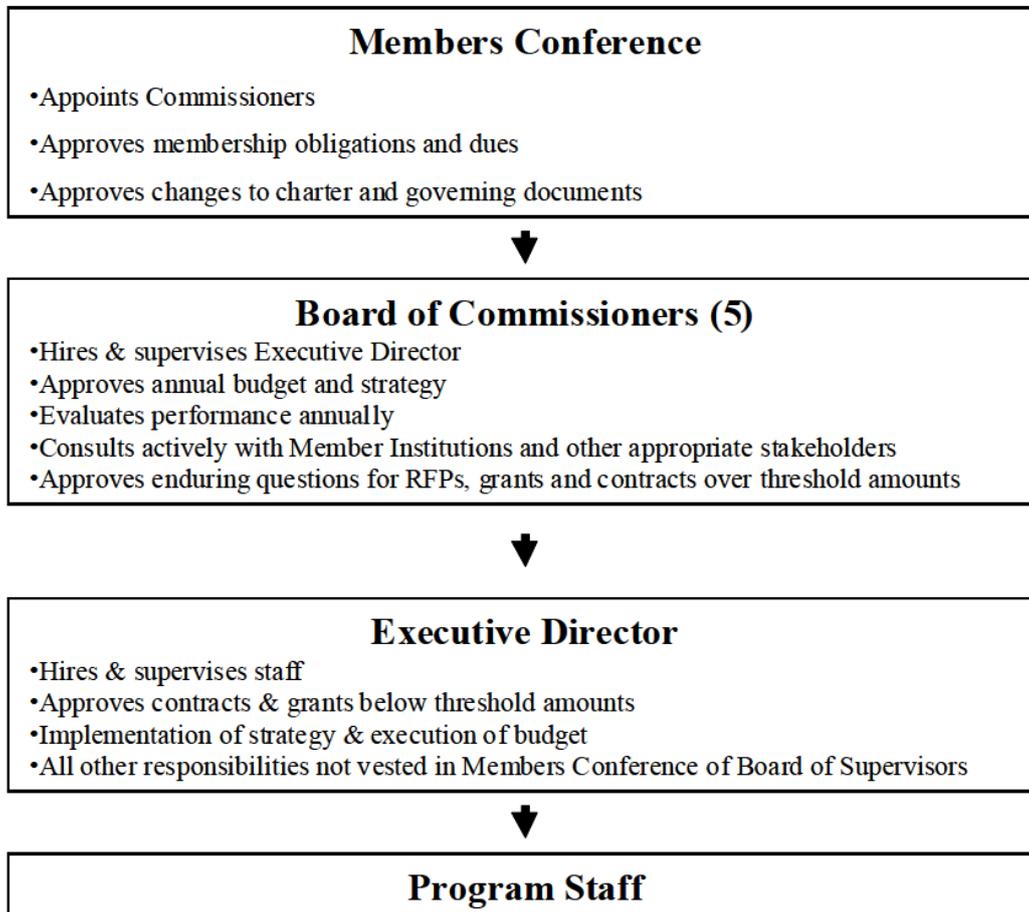
3.1.2 The Founding Document

122. The Founding Document for Establishing the International Initiative for Impact Evaluation (FD) was published on 25 June 2008.¹⁴ As described in Section 2, it sets out 3ie’s mission and operating principles, together with the activities it is expected to undertake. The FD also lays down the governance structure, summed up in Figure 4 below.

¹⁴ There is an earlier version from March 2007

Figure 4: 3ie Governance Structure

3ie Governance Structure



123. The Members’ Conference is to include both Member Institutions, i.e. those who pay a membership fee, and Associate Members from the IE practitioner community. Only full Members have a vote and the quorum is a simple majority of Member Institutions. Apart from paying the membership fee, Member Institutions are expected to be development agencies which implement at least \$1 million in social or economic development programmes annually.

124. The Board of Commissioners is appointed by the Members Conference to serve a three-year term. The Executive Director serves ex-officio, non-voting. Commissioners are expected to have a commitment the use of IE evidence in policy making. The objective is set that at least three should be from LMICs, two should have IE experience, and one should have financial expertise. One commissioner of international standing, to represent 3ie at the highest government levels, is also sought, and one to represent Civil Society ‘on behalf of lower-income individuals affected by development programs’.

125. When voting for the Board of Commissioners, Member Institutions are expected to divide into three groups: LMIC government agencies; multilateral and bilateral agencies; and non-profit philanthropic foundations and NGOs. A nominating committee of the Members Conference is to solicit nominees

and prepare a slate of candidates, with a nominee for Chairperson. The slate must include at least one nominee from each membership group.

126. As discussed below, these rather complex arrangements for electing the Board of Commissioners reflect a vision of the membership which has not been fully realised (though it is emerging). In other respects, the structure set out in Figure 4 is in place.
127. An Addendum to the FD sets out the proposed annual fees for membership of 3ie: \$20,000 for low-income countries and for private non-profit organisations; \$50,000 for middle income countries; and \$100,000 for multilateral and bilateral donors. These were minima, with the hope that all except LMICs would pay more.
128. At the time the FD was drafted, there was an expectation that Member Institutions might implement IEs. The Addendum states that 3ie will seek to ensure that grants awarded to any one Member would not exceed 25% of the total.
129. The Addendum delegates to the ED authority to approve contracts below \$250,000 and grants below \$1.5 million, subject to his following: expedited procedures (up to \$25,000); external review procedures; (up to \$100,000); or panel review procedures (up to \$1.5 million).
130. The arrangements set out in the Addendum were expected to be transitional, until the Members Conference and/or Board took action to replace them. In the event, it appears that the Addendum remained the only formal statement on these matters until the Board approved the Accounting Policies and Procedures Manual in April 2012. The new Manual makes no provision for the Board to authorise budgets and the only specific control is that the Board Chair will sign contracts over \$200,000 on the recommendation of the ED. The Manual does not mention approval of grant agreements, apparently leaving the Addendum provisions in that respect in place. As written, the Manual gives the Board of Commissioners no powers to review and approve budgets or major contracts and grants. In practice, the Board does review grant awards, as a recommended list, the annual Workplan and Budget, and larger contracts.

3.1.3 The Bylaws

131. 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. The following provisions are worth noting:

Membership

- Rights and Benefits of Members – For full Members, the only specific right/benefit is one vote on matters raised in member conferences. Representatives of Members and Associate Members may serve on committees, project teams, etc.
- Eligibility for membership is the decision of the Executive Director, unless the Board overturns his recommendation.
- There will be an Annual Members Conference and special Conferences can be called at the request of 25% of the Members.

Board of Commissioners

- As agreed in 2011, the Bylaws now provide for two Institutional donor seats on the Board, one for Governmental and one for non-governmental donors.
- Commissioners are now put into three classes: A, B, and C. After a transitional period, the aim is that one class will rotate off the board each year.

- Board action is permitted outside formal meetings, subject to the consent of all Commissioners in writing or electronically, such actions to be filed with the Minutes of the proceedings of the Commissioners.

Committees

- There are four standing committees of the Board: Executive, Audit and Finance, Governance and Program. There is also a Nominating Committee. Each Committee has a charter. Ad hoc committees are also allowed.
- 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.

Amendments

- The bylaws can be amended either by the Board or by a ‘majority of the Voting Members’. The only restriction on the Board’s power in this respect is that the Board may not amend anything decided by the Members for the first 12 months after the Members’ decision.

3.2 The Membership

132. The question of membership has been a difficult one for 3ie from the outset; particularly membership from developing countries (‘southern’ membership). The initial intention of membership was to promote IE broadly and to ensure that 3ie was guided by the needs of its stakeholder groups; ensuring that southern perspectives on the need for IEs were well represented and that this representation extended to 3ie’s governance structures and strategy. It was also hoped that the membership would bridge three distinct groups: donors, developing countries and IE practitioners. The first two were to make up 3ie’s institutional membership, while the practitioners would have a stake in 3ie’s work through non-voting associate membership.

3.2.1 Institutional Members

133. 3ie’s institutional membership increased from 12 in 2008 to 23 in 2011 and then fell slightly to 21 in 2012¹⁵. The membership is mixture of:

- philanthropic foundations	2
- bilateral donor agencies	10
- multilateral donor agencies	1
- developing country governments/ agencies	6
- INGOs	3

134. Only a core membership of two or three bilateral donor agencies and two philanthropic foundations contribute more than the minimum annual fee to 3ie’s core funding. Just under 85% comes from just two: one bilateral and one foundation.

135. The fee structure set out in the Founding Document has been slightly amended. LMIC members are no longer expected to pay a fee. Instead they are expected to provide in-kind support such as organising workshops. The rate for INGOs has been lowered from \$20,000 to \$10,000.

136. 3ie’s offer to members has three dimensions: a) an opportunity to demonstrate their commitment to aid effectiveness and impact evaluation; b) involvement in 3ie’s governance and oversight; and c) quality assurance and other technical services. Members are also offered a small advantage if they submit IE proposals for 3ie funding.

¹⁵ ED reports to the Board of Commissioners, May 2011 and April 2012

137. 3ie's External Liaison Officer (ELO), based in Delhi, is responsible for communication with the membership. The ELO's duties cover; organising the annual members' conference, reporting to members (donors) on the use of funds and meeting their reporting requirements, documenting the contacts 3ie has with different members, monitoring members' obligations and ensuring that any eligible studies and experts are documented. There does not appear to be a strategic dimension to this communication; it is more about ensuring compliance with various institutional issues.
138. The 2012 budget includes increased provision for supporting and expanding the membership and funding base, but the workplan merely states 'activities to support expanded Southern membership'. The main component seems to have been a workshop with southern stakeholders held in Rome at the time of the Membership Conference.
139. A 2011 review of members' needs¹⁶ noted that membership remained disappointingly low. Three distinct strategies were put forward:
- Stick with the current model: further developing 3ie's track record of providing quality IEs, will draw in southern members who want to participate in the process
 - Improve 3ie's communication to policymakers to drum up demand for membership, at the same time making the categories of membership more sophisticated
 - Recognise that southern members are not interested in being involved with 3ie's governance and identify non-membership based institutional models
140. There were two undercurrents in the survey results: first, the need to separate the issues of southern voice and governance and second, the need to more clearly define the benefits to members (including associate members). While there was support from Commissioners for having mechanisms that privilege Southern voice on strategic matters, they acknowledged that different approaches need to be considered.
141. As far as the Southern membership is concerned, 3ie's approach to ensuring policy impact via a particular mechanism (IEs) may in fact be getting in the way of developing a functional membership model. The main product on offer is impact evaluation. As the Policy Windows have shown, this can be a difficult one for institutional members to absorb, because of capacity weaknesses.
142. This suggests a fourth possible model: one where 3ie moves away from providing an IE with clearly defined expectations of immediate policy impact. Instead, it could offer a wider range of services relating to quality assurance and other forms of technical assistance which respond to local needs without the overriding imperative of demonstrating impact on policy.
143. The implications of this are that single demand-generation workshops in a country may not be enough. It may be more effective to set a programme to focus on a few countries each year, running several workshops in each country. This would help the national evaluation and policy communities to understand better what IE, and 3ie, have to offer. It would also overcome an issue raised by one interviewee, that the current strategy of granting membership to a country based on relationships with a single agency can privilege that agency unhelpfully; even create barriers to the involvement of other agencies. This strategy would imply broadening out the current demand-generation approach, even de-linking it from the immediate generation of Policy Window grants. The benefits of 3ie membership would come more from access to 3ie advice and services, and less from finding 'the impact evaluation that would serve as a gold standard in that country.'
144. Although bilateral donors are reasonably well represented in 3ie's membership, only a rather smaller the group is actively engaged. It is disappointing that there is only one multilateral. There is an increasing tendency among members in this group 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

¹⁶ D Fraser presentation to the Board of Commissioners, May 2011, Pretoria

community. This applies even to the two major funders, DFID and BMGF. As discussed elsewhere, this tends to weaken 3ie's core funding and strategic focus. There is a particular question about major donor agencies which use 3ie services, or even take 3ie grants, without paying even the minimum membership fee.

145. 3ie does not have a communication strategy for the membership. Annual Reports are designed for a wide public audience. Reporting to the Board of Commissioners, and through them to the Membership, could be improved. There is really little to suggest to members that 3ie considers them 'special' as a group. There is a case for PACO and AIES to focus on the membership as a whole, not as individual clients for services. The ELO's remit could also be broadened to include documenting and analysing 3ie's emerging understanding of members' needs for IE; communicating this between members and facilitating knowledge sharing and lesson learning – turning the role into something more strategically important to 3ie and the Board. These sorts of activities have been discussed in the context of developing a lesson-learning model similar to that operated by the DAC, instead of the C2E index. It would make strategic sense to incorporate this into a revised a membership strategy rather than an additional 'offer' from 3ie.
146. A stronger strategy would have two components:
 - 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 support a strategic, core-funded programme to deliver that public good.
147. 3ie is already working on both aspects, but they are unlikely to be successful without a strategic programme and dedicated resources to support it.
148. Developed world members would also need to agree to rebalance their contributions away from grants and towards the provision of technical expertise and, as in 3ie's original Theory of Change, advocacy for impact evaluation. But this advocacy would now be targeted on LMIC countries. This approach does risk slipping over into the capacity building role for which 3ie does not have the resources or the mandate; but a strict focus on quality standards for IEs, the importance of theory-based evaluation, and the use of mixed methods would clearly differentiate 3ie's effort from ordinary capacity building.

3.2.2 Associate Members

149. 3ie's Associate Membership is described as "*a community of development experts committed to improving lives through impact evaluation.*" The members are expected to demonstrate that they are capable of and actual do implement Impact Evaluations. Apart from 'membership in an emerging community of practice', benefits include: a 5% bonus score to AMs which apply for a 3ie grant; the opportunity to be considered for 3ie's expert panels; branding and promotion through publications, events, the 3ie website, etc.; and access to 3ie's expert roster and knowledge services. Being on the email list allows members to receive information on upcoming call for proposals, events, new research and publications. There are 102 organisations on the current AM list: 12 from Asia, five from Africa, four from Latin America, and 81 from OECD countries.
150. The Triple Line Stakeholder Survey has revealed a very lively community of evaluation practitioners, over 400 of them willing to give time to a survey on 3ie. How far this community gains from the 3ie Associate Membership status is unclear. The survey revealed that the networking aspects of 3ie's website were the least used. 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's treats its associates as special.
151. The Evaluation Team does not recommend that this should change. In fact, they would suggest it is no bad thing if 3ie maintains a certain distance from the practitioner community. If, as proposed, its

quality assurance role is reinforced that will increasingly put it in opposition to some practitioners. In the long term, 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, from study design through to final report and in selected cases, replication.

3.3 The Board of Commissioners

152. 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. Almost all those that the Evaluation Team have met agree that that the relationship was extremely difficult in the early stages. One Board meeting was described as ‘disastrous’. However, all were also agreed that matters are now much improved.
153. Apart from inexperience, there were important disagreements over both strategy and tactics, not just between the Board and the Executive but also within the Board itself. 3ie has made very important progress, with limited resources. Some strain was inevitable during the start-up of such a challenging programme.
154. Having attended the April 2012 Board meeting and reviewed the minutes of all previous meetings, the Evaluation Team believes that the Board has, in fact, worked rather better than many realise; certainly since the concerted effort to address institutional issues in mid-2011. Where there were difficulties the Board and ED sought to address them, either directly or by taking advice from consultants. This has been a joint effort, with some members of the board taking an active role, especially in sub-committee. This may have been a factor in the establishment of five separate sub-committees. This was discussed in the seventh Board meeting. The Evaluation Team agrees with those Commissioners who argued that the sub-committee structure risked the Board dealing with matters that should be left to the Executive. It should be the Executive which does all preparatory work, while the sub-committee’s task is to review the final proposal and ensure that it is ready to be presented to the full Board. If this happened, it should be possible to do without so many standing committees. When needed, an ad-hoc committee would be sufficient.
155. However, it is clear that until the recent creation of three new Deputy Director positions, the Executive has lacked the resources to address all the issues the Board raised. From the minutes, some issues have rolled-over unresolved, from one meeting to the next. And some Board requests for a paper on an issue have been answered with a few paragraphs in the ED’s report instead of a full paper. For that reason, greater Board involvement may have been inevitable. 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

156. A 3ie Commissioner described the Founding Document’s vision to the Evaluation Team as an organisation which is ‘donor-funded, southern driven’. This was always going to be an ambitious vision to realise and it is clear that 3ie is still some way short of it. Although there are a number of practical weaknesses which can be addressed, it has to be recognised that this vision will not be realised without a targeted effort, and a budget to match.
157. At present, the Membership as a group is not strong enough or cohesive enough to hold 3ie to account robustly, with oversight of implementation and the final word on strategy. This was not perhaps its role. The Governance Structure shown in Figure 4 only gives the Members Conference powers at the very highest level, with no role in either strategy or performance. This in itself may have weakened ownership and contributed to a tendency for the membership to deal with 3ie as individuals, not as a constituency or even set of constituencies as originally hoped. Even within this limited remit, one critical function has since been transferred from the Members Conference to the Board of

Commissioners. The Founding Document expected the Nominating Committee to be part of the Members Conference, but the bylaws now have it subsidiary to the Board. However it is not clear what this subsidiarity means, since the Board only appoints 2 out of the 5 members.

158. Nevertheless, 3ie management do present a full report to the Annual Members conference, and there is considerable discussion of strategic issues and some operational matters as well. This process shares the weakness discussed in Section 3.5, the lack of a planning, budgeting and reporting process which is fully effective for oversight and control. Nevertheless, members have had a chance to hear challenging reports, such as the Rapid Institutional Review.
159. The new 3ie Board of Commissioners takes office in October 2012. At the Members Conference which approved the new Board, a concern was expressed about the narrowness of the representation. There are four members from OECD countries, two each from the US and the UK; four from middle income countries, two of them from Mexico; and, two from low income countries. A majority of the Southern representatives are development specialists, trained internationally, some with long experience in international organisations. While most of the requirements set out in the founding document have been met, there is a concentration of IE expertise. Even the representative of Civil Society ‘on behalf of lower-income individuals affected by development programs’, is an evaluation specialist from a US-based INGO. Overall the concern expressed may have some force. The Southern voice has perhaps a rather international accent.
160. 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 by donor representatives, 3ie staff, associate members and even consultants. Even so, only four LMICs were represented, compared with nine northern organisations. It is also worth noting that the 3rd meeting did not have a quorum of a simple majority of members, northern and southern, and the 4th meeting was only just quorate.
161. 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. Had there been strong accountability to the Members Conference, it is possible the funders would not have felt the need to press this issue. However, it is only proper that their critical interest is fully represented. The Evaluation Team believes it far better that this is done in a transparent way, and through open discussion in Board meetings. The alternative is for there to be separate lines of accountability to individual funders. This should be avoided if at all possible. Similarly, it is far better for the ‘donor funders’ to engage openly with 3ie’s ‘Southern Drivers’ in the Board forum, than it would be to try to establish some other channel for this.
162. 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. It might, however, help to reinforce the formal role of the Membership Conference, emphasising that it is a business meeting at which the Membership holds the Board and 3ie accountable. Tighter control over attendance would help with this. The presence of Associate Members, consultants and other stakeholders can create an impression that the Conference is a workshop, not a business meeting. It also tends to worsen the imbalance between south and north. Developing a membership strategy (see section 7) would offer a useful opportunity to look at the role of the Membership Conference.

4 Stakeholder Consultation

163. As part of the Process Evaluation, three groups of 3ie stakeholders were consulted through an on-line survey:
- The **Contract Negotiator (CN)** survey - addressed to the contract officers on grants awarded under the first three open windows. This part of the survey focused on the financial management of the grants and the 3ie grant management process.
 - A **Principal Investigator (PI)** survey - addressed to the lead PI on each grant and focused on the technical management of the grants, in particular the implementation experience and policy uptake.
 - A **Stakeholder survey**. This survey was more general, casting a wide net to capture the views of four categories of stakeholder: developing country policymakers, donor organisations, implementing agents and researchers/evaluation practitioners. The Evaluation Team considered ways to reach these groups separately, but could not find a workable framework. The most practical approach seemed to be to contact the widest possible community, and ask respondents to identify which category they felt they belonged to.
164. All three were e-surveys via the internet using Survey Monkey, with the first link sent out by email¹⁷ 3ie provided email lists for the CN and PI surveys. The Stakeholder survey was distributed through 15 listservs, which served a wide range of international impact evaluation users and stakeholders. This drew responses from a large community of over 500 IE stakeholders. Table 16 summarises the response to each survey. Because of time constraints it was not possible to fully pilot the survey, which resulted in a few unfortunate omissions and poorly worded questions.
165. All three surveys followed a standard format, in most cases asking respondents to identify which of four or five descriptions best matched their experience or opinion. In reading what follows, this should be borne in mind. For brevity the report states, for example, that “most found 3ie’s requirements were ‘reasonable’”. This should be read to mean that most selected the rating ‘reasonable’. Comments and other verbatim responses are noted where appropriate.

Table 16: Stakeholder Response Numbers

Survey	Number of responses	OW Grants covered	Surveys completed in full
Contract Negotiators	22	30	17
Principal Investigators	37	39	36
Stakeholder Survey	510	22	409

4.1 Contract Negotiators

166. The CN Survey concentrated on grantee experience with 3ie’s grant management procedures. To avoid duplication, aspects already covered in 3ie’s own Applicant Surveys were not repeated (See Section 2.5). 22 CNs responded to the survey: five from OW1, 12 from OW2 and five from OW3. As

¹⁷ Links to the three surveys can be found at <http://www.surveymonkey.com/s/3iePI>, /3ieCN and /3ie for the Principal Investigator, Contract Negotiator and Stakeholder survey respectively.

four of them hold more than one grant,¹⁸ they represent well over a third of 3ie's 72 Open Window grantees. With 10 respondents from the USA and three from Europe, and others from Chile, China, Mexico, India and S. Africa, most were based in high or middle income countries.

4.1.1 Grantee Organisations

167. The first section sought a description of the grantee organisation's status and experience. The following sums up the results:

- A majority of grantees are established institutions with large portfolios of 10 or more current research grants. Around half of them have over 40 current grants.
- Correspondingly, 3ie funds make up only a small proportion of their research budgets: less than 10% for more than half of the respondents.
- If they had not received 3ie funding, most would have looked elsewhere for funding: 12 out of 18. Only two said they would not have done the study without 3ie funds.
- The majority (16/21) were experienced and had managed at least 10 previous studies using experimental/quasi experimental research designs.
- Well over half (13/21) were specialist institutions for which experimental/quasi experimental research makes up more than 50% of their work. Of the other organisations, the main activities were in applied research, policy impact studies, rapid appraisals and econometric modelling.

168. To sum up, the majority of 3ie grants are managed by well-established institutions which specialise in the kind of research 3ie aims to support. For these institutions, 3ie funding seems to be useful but not critical to their operations.

4.1.2 The Contracting Process

169. CN's were asked to describe their experience in completing their agreements with 3ie. About half of those who responded described the amendments 3ie sought in their budgets as 'substantial' or 'significant'. Although most described 3ie's requirements as 'reasonable', a significant minority (7/21) felt they were 'excessive'. Large majorities – 75% and over - described 3ie's instructions to grant applicants as at least 'adequate' and reported that 3ie had completed budget negotiations 'with minor delays' or better.

170. Some specific difficulties were reported. The most important were the retroactive application of new reporting requirements, such as Policy Influence Plans, and the difficulties some US universities have with the 3ie-GDN grant agreement provisions on audit, liability and confidentiality. Given that GDN's main activity is providing grants to LMIC universities it is difficult to see the reason for this. However the evaluation team have not investigated further.

4.1.3 Financial Arrangements

171. 3ie pays grants in tranches. After a relatively large first tranche – 40%+ on signature of the grant agreement - subsequent tranches are paid against receipt of specified deliverables, such as detailed survey designs, draft questionnaires, interim batches of data as well as progress reports. Of the CN respondents, most grants have three to five tranches, of which they have been paid between one and three so far. Tranche intervals range from 3 to 22 months but for most grants it is 12 months. Almost all grantees thought these payment intervals are 'about right'.

172. 3ie follows GDN guidelines on financial reporting and GDN staff monitor budgets and make the tranche payments on behalf of 3ie. A minority (3/17) found the financial reporting requirements 'too

¹⁸ One held 2, two held 3, and one held 6)

detailed and bureaucratic’, but the majority found this was also ‘about right’. With one exception, 3ie tranches were paid ‘promptly’ or with only ‘minor delays’.

173. Of 17 who answered the question, eight CNs had requested a Letter of Variation to their 3ie grant, for a time extension or budget variation. Reported reasons included changes to the scope of the work, the research design, or the deliverables schedule; adding a qualitative component to a study; and adding a second round of program intervention to a study. Others were to allow for delays in completing deliverables.

4.1.4 Grant Funding

174. The survey revealed that 3ie is not the sole or even the majority funder for most of the grants it supports. Of 17 responses, 3ie was only sole funder for five grants. Of the 12 with co-finance, seven received more than half their funding from other sources. With only two exceptions, these co-financing arrangements were not supported by an MoU between the different funders. With only one exception, grantees report to each of their co-funders separately, ‘with different timescales and deliverables for each donor’.

4.1.5 CN Assessment of 3ie

175. CNs were asked to rate 3ie’s financial and contract arrangements, bearing in mind their experience with other funders. 11 of 18 chose the rating ‘firm but fair’ but there was a significant minority (6) which selected ‘excessively bureaucratic’. Only one saw 3ie as ‘light touch’ relative to other funders. 3ie scored well on financial arrangements, with 13/17 saying it has managed financial aspects ‘efficiently’, and the rest selecting ‘efficiently, with occasional hiccups.’ There was one comment that things had ‘started rough during negotiations but improved greatly during implementation’ and one that templates in Excel, not Word, and clearer instructions would be an improvement.¹⁹

4.1.6 Conclusion

176. Overall, CNs seem happy with the way 3ie is managing its grants. A solid majority give satisfactory to good scores in almost all respects. There is a minority, perhaps a third, which feels 3ie procedures are too bureaucratic. On balance, that may not be a bad thing. One final CN comment sums up the balance: *“It is easily understandable how 3ie is seeking to make this process efficient and cost effective for itself, but it needs to keep in mind that it needs to be careful about the burden it is placing on its non-profit recipients.”* CN responses seem to indicate that 3ie may have got the balance ‘about right’.
177. One clear point emerges from this part of the survey: that most grantees are well established, experienced specialist institutions and they are well funded. 3ie funding is a useful addition to their resources but their research is not dependent on that funding.

4.2 Principal Investigators

178. The PI survey concentrated on grantee experience with the technical review process, project implementation and policy uptake. 37 PIs responded, three from OW1, 14 from OW2 and 22 from OW3. With two PIs working on more than one grant, respondents represent over half of 3ie’s 72 open window grants. 25 of the 37 were from the US, and three from Europe. The remainder were mainly from middle income countries: Chile, China, India and S. Africa.

¹⁹ 3ie now provides a detailed Excel template.

4.2.1 Grantee information

179. The first section sought a description of the grantee PI's status and experience. The following sums up the results:

- The majority of PIs were professors or assistant professors. (25/37) The others were scientists, economists or directors. 22 were from US Universities, and three were from the World Bank.
- A majority (30/37) spent between 10 and 30 per cent of their time on their 3ie research project, with no-one spending less than this, and only two spending more than half their time on 3ie work. Most (29/35) spent the majority of their time on other research, with a small minority spending it on teaching or management.
- Most PIs were experienced in leading experimental/quasi experimental research studies in developing countries, with 28/37 having led at least one before, and 11 having led at least 5 previous studies. However, there were three PIs for whom this was their first time leading such a study.
- A large majority of PIs described their knowledge of the study country study as reasonable (8) or extensive (21). Seven said they relied on the extensive knowledge of a co-PI and one that he relied on the experience of the implementing agency.
- The main research interests of PIs varied, with the most popular being health (10), and agriculture (8), followed by governance and education.

180. To summarise, most PIs are experienced researchers, who spend a significant portion of their time on 3ie work, as part of other research in a variety of fields.

4.2.2 Study information

181. The next section sought the PIs description of where the 3ie grant fitted into their other work:

- A majority of studies were either a direct follow up (11) or an extension of previous work (17), with less than a quarter being in a new area of work for the PI (9).
- Without 3ie funding, a majority of PIs responded that they would have either found different funding for the same research (14) or conducted the research on a smaller scale (12). A significant minority said they would have done a different piece of research (10). Only one said he would have had to use the time for non-research activities
- A majority of PIs (21/37) took the lead in identifying the research issue and putting the team together and then approaching the IA (implementing agency) to do the study, with a minority being invited by another researcher (12/37). Only four had been asked to do the IE by the IA.
- All PIs thought that their study was either a significant extension of current knowledge (20), or a major step into new knowledge (17). Similarly, a majority (23) thought that it will influence important current policy debates, or even lead to major policy change (9). Only three PIs gave the lower rating of 'a useful contribution to policy under specific circumstances.'

182. The fact that most grants spring from earlier work and that it was the PI, or another researcher, who took the lead in the identifying the research issue, does suggest that most 3ie grants can be described as 'researcher-led' or 'supply-led'.

4.2.3 The Technical Process

183. PIs were asked to describe their experience in completing their agreements with 3ie. The majority rated the revisions 3ie sought in their Comments for Response as 'minor'. Only seven (/37)

respondents described them as significant, and one as substantial. The vast majority described 3ie's requirements as 'reasonable', and only three felt they were 'excessive'. A majority also felt 3ie's instructions to applicants were at least 'adequate'.

184. 3ie provides grant applicants with a very complete set of background material: Instructions for Proposals; Principles of Impact Evaluation; Impact Evaluation Practice; Human Subject Guidelines; Indirect Cost Guidelines and FAQs. The majority found these useful. However, between 15% and 25% either did not read the technical guidelines on IE principles and practice, or did not find them useful. This may indicate a feeling that researchers do not need to be told these things. One PI claimed not to have read the Instructions for Proposals, perhaps indicating he was not closely engaged with the grant.
185. Twelve respondents reported that 3ie had completed the technical review process 'promptly', seventeen 'with minor delays'. Some specific misunderstandings were mentioned which had caused delays: concerning the study design, and with obtaining secondary documentation, such as communication plans.
186. Although the majority thought the technical deliveries specified by 3ie for grant monitoring to be 'about right', a sizeable minority (14/36) thought them too detailed and bureaucratic. Unsurprisingly perhaps, no one thought that the technical deliverables specified were too loosely specified.
187. 3ie reviewers' comments were seen as being 'sometimes useful and sometimes not' by well over half of the respondents, but 14 of them rated them as 'technically sound and helpful'. Complaints about the reviewers' comments were that they focussed on minor points, that approaches were misunderstood, or that they requested changes which were useful for clarity of communication, but not for the study design. Asked if 3ie staff respond quickly to queries and provide good support, 30 PIs said 'Yes'. Some reported problems, especially when staff changed. However, even the most negative remark ended on a positive note: *"It has sometimes taken several weeks, and multiple follow-ups, to get a response from 3ie staff. I think this has gotten better over time, though."*
188. As a whole, 3ie's technical review and support processes were seen as 'firm but fair' by two thirds of PIs, with seven respondents who found them 'excessively bureaucratic' compared with other funders. On the other hand, four found 3ie 'light touch'. Of those finding the processes too bureaucratic, one mentioned the PIPs as being counterproductive, one that it involved unnecessary repetition, and one that it takes too long before tranche disbursement. The most negative remark was: *"3ie asks for more things and has more forms than any other funder I work with. Part of it is adding in a whole piece on policy influence and that at each stage you get questions again on your original application."*
189. Overall, it would seem that 3ie's technical processes are indeed 'about right': providing a steady and technically capable challenge to grantees; a challenge to which a majority are happy to respond. If there were no complaints, it would suggest that 3ie had become a 'light', if not an 'easy touch'. (As is discussed at some length in Chapters 6 and 7, the Evaluation Team believe to be of the greatest importance that 3ie maintain, or even strengthen their technical oversight.)

4.2.4 Study implementation and Surveying.

190. 11 of the 37 PIs reported substantial changes to the proposed study design during implementation. Reasons included political turbulence (2), extensions or delays to the study (5), moving the study (1) or changing PIs (3). These changes were reported to 3ie, and in all but one case they handled the matter 'flexibly to achieve a solution'.
191. Although most PIs have an MoU with the agency implementing the intervention (IA), 11 did not. In the vast majority of cases the management and supervision of the selection of clusters for sampling, as well as of individual households was done by the research team under the guidance of the PIs. However, in five cases this was done by the IA or by a survey company or third party organisation.

192. Most surveys are implemented by ‘contracted staff trained and led by the research team’, 12 used a contracted survey company or third party organisation. In four cases, it was the staff of the IA who implemented the survey.
193. Almost all respondents were either ‘satisfied’ or ‘very satisfied’ with the survey implementation. Only one, whose survey was implemented by a third party agency, was ‘dissatisfied’. However, the comment “Glad it’s over” by one PI may indicate that things were not always easy.
194. 11 PIs reported that they had had to ‘compromise their survey design for practical reasons’. None of them thought that the changes were serious: either ‘not important’ (3) or ‘manageable by statistical techniques’ (8).
195. On this evidence, the PIs have faced few problems implementing their studies. This does not match the Evaluation Team’s findings from a detailed review of implementation on a sample of grants. It may indicate that some PIs do not give sufficient attention to implementation. (See Section 6.3.3)

Policy Influence

196. Asked if they felt they understood what policy influence meant when submitting their proposal, a quarter said they already had ‘a good understanding’. However, a majority of PIs felt ‘3ie provided useful information to clarify what they meant’. Although PIPs only became mandatory for OW3, all but one of the PIs responding to the survey had done a PIP.
197. Almost exactly half of PIs rated 3ie’s PIP requirements ‘excessive’, and half ‘reasonable’. This split was highlighted by the way one very negative comment was balanced by one very positive: “Please note some of us work not to publish papers but to influence policy, we don’t benefit from guidance of this kind and find it tedious”, compared with “PIP rules are a stretch, but a good one, the guidance is very helpful, as are specific suggestions from 3ie staff.” Two more nuanced comments suggested a need for flexibility, especially about timing: PIP requirements are excessive at early stages, but should be emphasised later.
198. This balance of views is reflected in other responses. Most found the process of developing a PIP at least ‘sort of’ helpful (22/37), but six did not. The stakeholder engagement plan was seen as the most difficult section to fill in, followed by the stakeholder and policymaker analysis and the evaluation and learning sections. When submitted, over half of PIPs required minor revisions, and a quarter significant ones. However, three quarters of respondents found the 3ie feedback on PIPs helpful, and timely.
199. A key question concerned PI views on the advocacy role implicit in the PIP. Just over half reported they are comfortable with that role, and believe they can fulfil it well, and three are comfortable with the role but need training and support from 3ie. Against that, over a third (14) PIs are uncomfortable with an advocacy role. Two of them feel it compromises PI neutrality. The comments revealed strong views. One PI commented that the PIP was actually detrimental to policy influence, because getting letters for each disbursement can kill political capital, and researchers advocating for certain policies can be unhelpful, very time consuming or not appropriate until results are out. Others felt that advocacy was not a comparative advantage of researchers, and should thus be done in conjunction with other agencies such as JPAL. A useful role for 3ie might be to link researchers with policy-oriented organisations.

Thinking on policy influence

200. PIs were asked how their views around policy influence had changed over the life of the projects. 15 replied that their views had not really changed. Of the others, 14 had found the information provided by 3ie useful.
201. Seven respondents had attended a 3ie policy influence clinic. Six of them would recommend the clinics to a colleague. One other PI who had not attended said he would now based on the recommendation of a colleague who did.

202. PIs were asked if they had carried ideas from their work on policy influence with 3ie into their other work. A majority did not answer this. Of those that did, only seven reported a suggested a significant impact: ‘a review of their approach to research communication and policy engagement’.

203. When given the opportunity to make a final free response at the end of the survey, five PIs came back to the question of PIPs, reinforcing the strength of opinion on this:

A. *“3ie’s policy influence plan seems to ignore the complexities.” ... “It is also highly unrealistic to ask researcher to list things like: where the paper will published, how many newspaper citations, how many policy makers will use the results, etc. You cannot promise how other people use your work...you can only try to do good research and try to make the results known to the right people.”*

B. *“Getting a letter from [Government officials] on *anything* -- even an email for 3ie purposes -- requires a huge effort as it needs to be cleared by various bureaucratic hurdles. It costs us substantial political capital to get these letters, political capital that would be much better spent advancing policy discussion and working hand in hand with them on new policy evaluations. Ultimately these letters I think hurts our relationship with government and makes it harder for us to effectively communicate, collaborate, and work on important policy questions”*

C. *“I am very uncomfortable with the policy influence plan, because it requires active engagement and attempts to influence policymakers before there are reliable results from the analysis conducted. It seems to me that this could be disastrous.”*

4.2.5 Conclusions

204. Overall PIs are experienced academics who are also satisfied with the 3ie guidance and grant management processes, as well as study and survey implementation. There was a minority which found 3ie reporting requirements to be burdensome. *“The work asked by 3ie from PIs is way too much and definitely not commensurate with the size of the grant (i.e I received only \$100K but I feel that I was reporting as if it was a \$1M grant).”* A useful conclusion from this comment might be that 3ie should not consider very small grants.

205. Almost half are dissatisfied with the PIP process for a variety of reasons. The requirement for letters from stakeholders is a particular issue. Some PIs find them a waste of time, or even counterproductive. A number of PIs feel the advocacy role implicit in the PIP may be inappropriate, especially before the study is complete and the policy implications of its results have been considered.

4.3 Stakeholder Survey

206. The stakeholder survey was distributed to evaluation practitioners, bilateral and multilateral agencies, donors and developing country practitioners. It was sent to over 5,000 email addresses though 17 different email lists. 408 respondents fully completed the questionnaire, and a further 101 partially completed it. Replies came from over 70 different countries: a quarter each from Europe, N. America and Africa, with the remainder split between Asia and South America. The nature of the community surveyed can be summed up as follows:

- Respondents were highly experienced in international development. 73% reported more than 5 years of experience in the sector.
- Just under half described themselves as evaluation practitioners (210 respondents), and a quarter as researchers (143 respondents). The remaining quarter were divided between Government agency policymakers (27), development programme managers (41), users of IE (34) and the donor community (17).

- By organisation, 159 came from Universities/Research institutes, 104 from consultancy companies, and 70 from INGOs. There were 42 responses from Developing Country government agencies. Between bilateral (17) and multilateral agencies (26), and foundations (8), the donor community was also well represented.
- Of the 408 completed surveys, 24 were from 3ie member organisations, 23 from associate member organisations, and 17 from organisations who have partnered with 3ie. The remaining 84% were from organisations not affiliated with 3ie.
- Respondents came from a wide range of evaluation organisations (over 50 different evaluation networks and associations were mentioned). More than 50 respondents each were members of AeA, IDEAS, and My MandE.

207. Overall, the survey drew a sizeable response from all groups with a major interest in Impact Evaluation.

4.3.1 Interest in Evaluation and Experience with 3ie

208. Although it does reflect the fact that the main email lists used for the survey were provided by 3ie, an impressive 91% had heard of 3ie before: about a third each from colleagues, through an evaluation network, or direct from the 3ie mailing list or website.
209. 83% of respondents had a general interest in impact evaluation, and half a specific interest in 3ie. Only 36 respondents were involved with 3ie directly: as grant holders, reviewers, staff, speakers at 3ie events, or board members.
210. Reasons for being interested in 3ie included: a wish to keep up with best practice (70%), to improve the ways of doing evaluation (62%), to find more robust ways to test impacts (51%) and to improve programme implementation (40%). A core group, nearly a third, have been involved in evaluation debates since before 3ie was founded.
211. 135 respondents had attended a 3ie conference or other event.²⁰ Feedback from 3ie events was largely positive, with respondents reporting that the conferences were well organised, very interactive and interesting, with lively debate and experienced speakers. Some felt that the London seminar series was at times not well attended, and or that other talks were too short, or participants too many. A couple of respondents felt that the events lacked scientific rigour and should be open to more supplementary techniques as opposed to just experimental/quasi experimental ones, although others praised the range of methods presented, with sessions given on when different methods have most value added. Further comments came from participants unable to attend events due to lack of funds, and seven expressed an interest to attend future 3ie events.
212. Asked which 3ie publications they had read, a larger than usual proportion skipped the question, and between 40 and 50% of those who replied had not read them. Overall, between a quarter and a third of those surveyed found 3ie's published material a useful resource. Both on numbers reading and the best scores for 'very useful', 3ie's Systematic Reviews and Working Paper series rated highest.

²⁰ 3ie has run separate surveys of those attending events so the Process Evaluation did not cover this in detail. See Section 2.5.

Table 17: Usefulness of 3ie Publications

	Read	Very useful	Useful	Not useful
3ie Systematic Review	56%	27%	26%	3%
Journal of Development Effectiveness	46%	18%	26%	2%
3ie Working Paper Series	62%	25%	36%	1%
3ie Evidence Matters Brief	46%	16%	27%	4%
3ie Brief	48%	17%	31%	1%

Note: Per cent of those answering the question.

213. Asked the same question about the different components of the 3ie website, the response rate was slightly higher. The Used/Not Used split varied widely. 86% have used the site as a whole, and almost all of them found it useful. The IE and SR databases were used by 50% and the other four elements by only a quarter of those who responded. It is possible that the grant funding parts of the website are what respondents found most useful. Unfortunately this question was not in the survey.

Table 18: Usefulness of 3ie Web Resources

	Used	Very useful	Useful	Not useful
3ie Website	86%	34%	49%	2%
3ie Blog	27%	6%	16%	5%
3ie Forum	22%	5%	13%	4%
3ie Social Network	26%	7%	12%	6%
3ie Expert Roster	23%	8%	8%	7%
3ie IE Database	56%	21%	32%	3%
3ie SR Database	51%	21%	26%	4%

214. To conclude this section of the survey, all respondents were then asked to say how useful they found 3ie’s work for their own work. Figure 5 sums up the results. Despite the varied responses to the earlier question on the different 3ie publications, as a whole they were rated highest, with 70% describing them as useful/very useful, and almost all the rest thinking that they could be useful for work in the future. Interestingly, Systematic Reviews rated rather higher than IE grants for being currently useful, although the position was reversed for the IE and SR databases. The ratings were lowest for the 3ie workshops and, especially, the expert roster, which barely a quarter of those responding to this section of the survey found currently useful.

215. To sum up, it is clear that 3ie has established a very strong position within the stakeholder community consulted. Although there must be bias from the fact that respondents are self-selected, there is very high awareness of 3ie and its activities. Almost all of these are rated very highly. The Evaluation Team has not had time to deepen the analysis by, for example, seeing how much ratings differ between the various stakeholder groups. They would be happy to share the data with 3ie for that analysis to be done.²¹

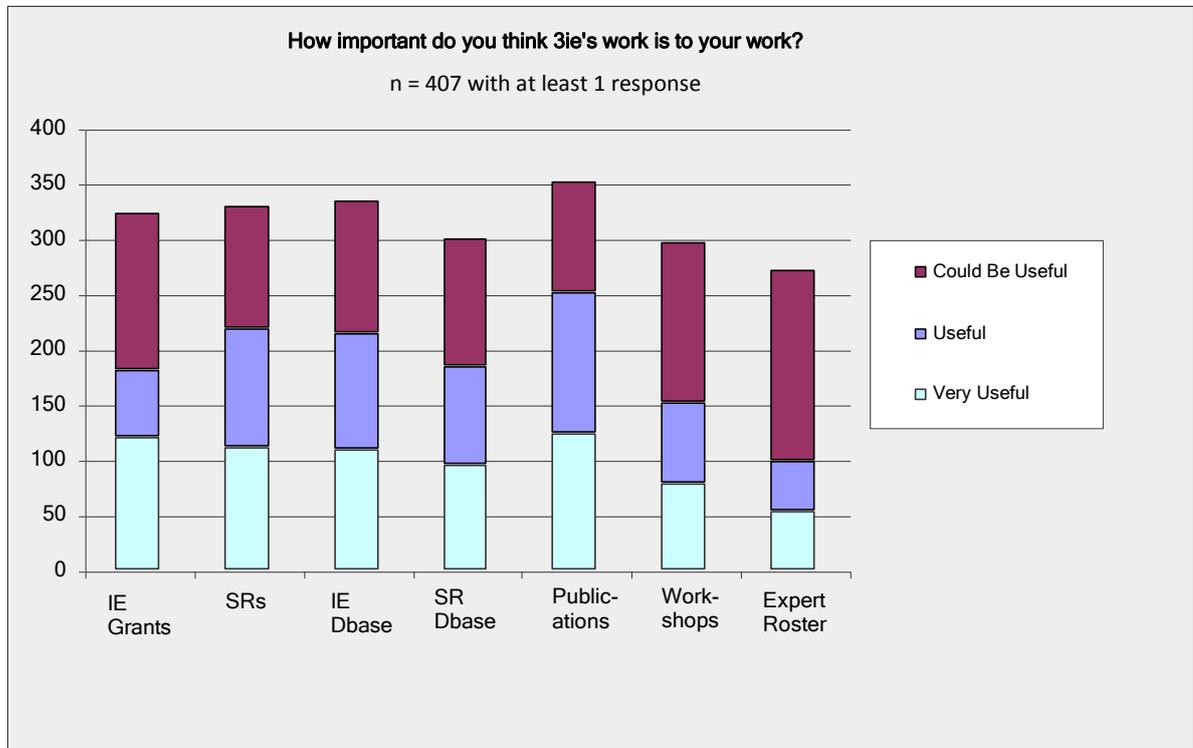
VIEWES OF DIFFERENT STAKEHOLDER GROUPS

216. From this point on in the Survey, stakeholders were asked to select different sections depending on their role and interest in Impact Evaluation. The vast majority of respondents described themselves as evaluation practitioners or researchers (330/72%). Of the remainder, 66 were development programme managers, or implementers, 21 Government or government agency policymakers in developing countries, and 22 representatives of funders or donors. Each group was asked to answer

²¹ Triple Line did promise respondents that their replies would be confidential, so it would be necessary to anonymise the data to do this.

a section of the questionnaire tailored to their role, with the option of answering multiple sections if they felt their role overlapped.

Figure 5: 3ie's Importance to Stakeholders



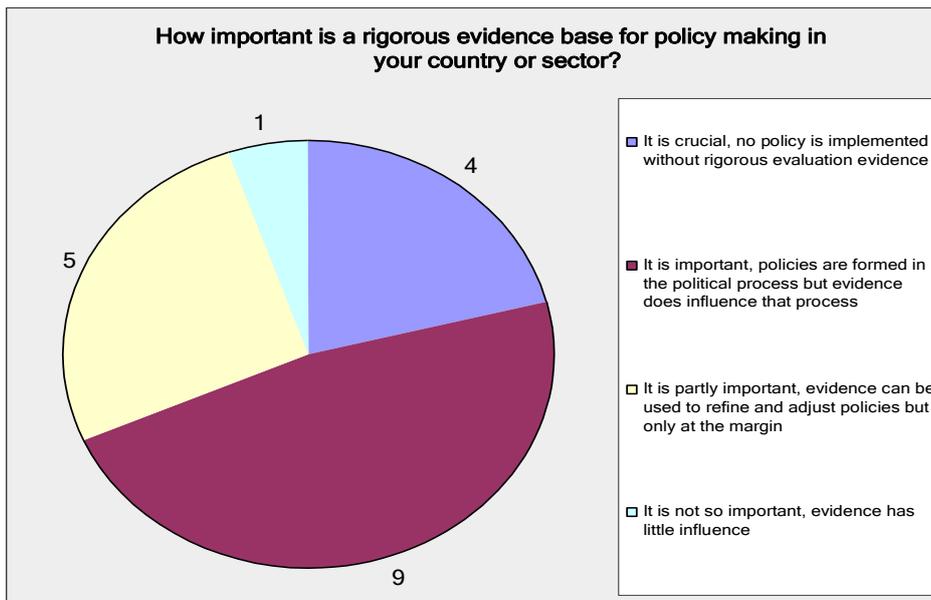
1. GOVERNMENT OR GOVERNMENT AGENCY POLICYMAKERS IN DEVELOPING COUNTRIES

217. Only 17 respondents answered all the questions for developing country policymakers. By sector, four were specialists in Agriculture, and four in Governance. The others were split between 9 other categories including: environment, science, IT, justice and finance. 10 developing countries were represented including: India, Jordan, Nigeria, and Pakistan, with two or more respondents each, and Columbia, Indonesia, Peru, Sierra Leone, Romania and Uganda.

Policy formation

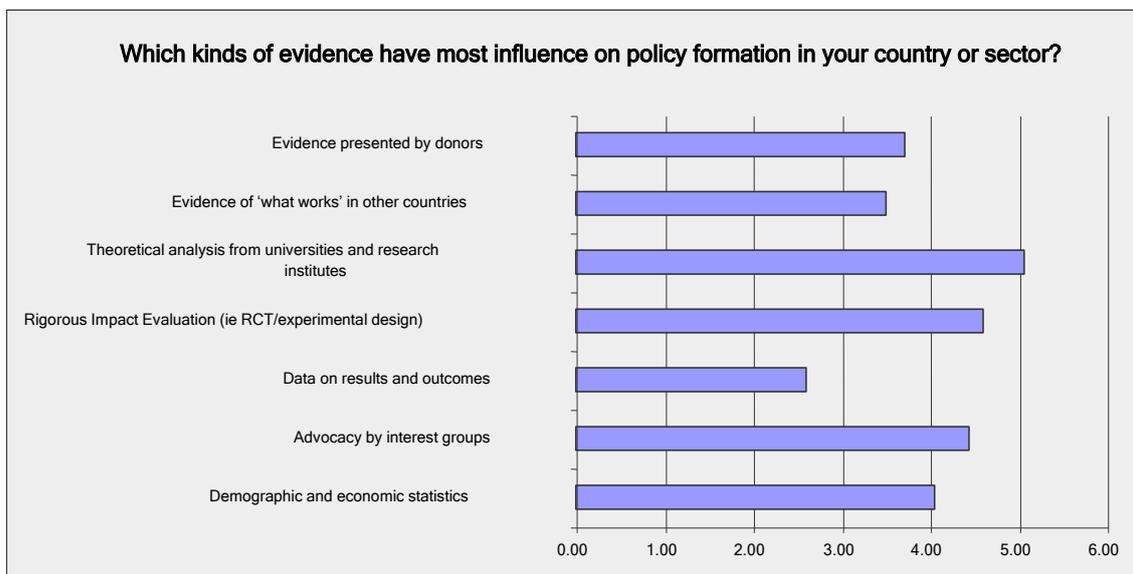
218. DC policymakers were asked how important a rigorous evidence base is for policymaking. Figure 6 summarises the response, showing a realistic view: that evidence may have influence but politics will always come first.

Figure 6: Influence of Evidence in DC Policymaking



219. Respondents were then asked to rank what kind of data is most important for policy formation. Figure 7 shows the results by average rank, where 1 is highest.²² It shows how data on results and outcomes is considered by some way the most important. RCT-style evaluations and theoretical research rank lowest.

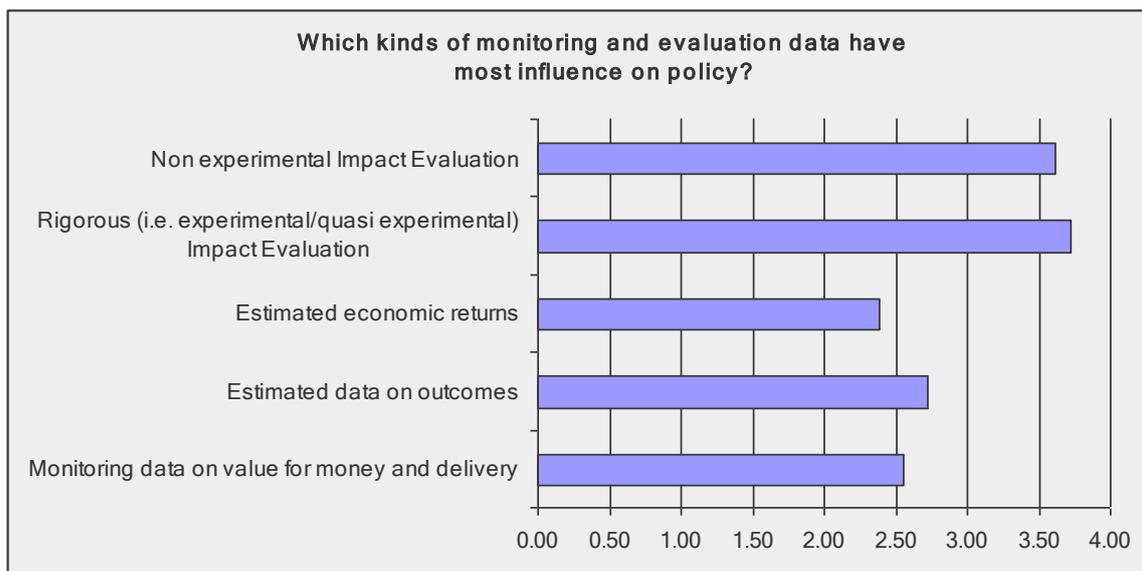
Figure 7: Influence of Evidence on Policy Formation?



²² This is the standard for the figures which follow.

220. Asked about other kinds of evidence, two respondents mentioned the political situation, with an example of how a new form of health insurance helped an Indian party win a state election. Two other comments were ‘donors’ views’ and ‘feedback from community fora’.
221. Policymakers showed a clear preference for traditional M&E techniques as opposed to Impact Evaluation, whether experimental or not. Estimated economic returns, data on outcomes and monitoring value for money and delivery were all rated clearly higher. Figure 8 summarises the response. There is even a slight preference for non-experimental over experimental Impact Evaluation.

Figure 8: Policymakers’ M&E Preferences



Evaluation funding and Use

222. Although multiple responses make the result unclear, about half of policymakers reported that their Government evaluates its programmes through a central evaluation unit and half through an evaluation department in each Ministry, Department or Agency. Most respondents (16/19) reported that their Government does conduct Impact Evaluations. Over half responded that most or all evaluations are reliant on donor support. Only three reported that all IEs are wholly funded from departmental budgets.
223. When asked when how Impact evaluation evidence has been used in the past, most replied that it had been used at least ‘a little’ in developing long term policies and strategies, and in assessing the effects of specific policies or programmes. It had been used less in deciding budget allocations, with six of 19 saying it was not used at all for that purpose.
224. A final question asked developing country policymakers views on the tension between the need for evidence, and the time it takes to ensure that conclusions are robust. Just over half (10/18) selected the answer “*It’s very important to have the evidence when I need it so I generally need to rely on rapid evaluation techniques rather than full impact evaluations*”. However, the other eight opted for answers underlining the need for rigour and the risk of using incorrect data. One written comment suggested it depended on context but concluded: “The worst approach is to wait until the perfection is reached!”

Summary of developing country policymaker responses:

225. Despite the small number of respondents, this section of the survey paints a clear picture of the practical concerns which are foremost for developing country policymakers. They are aware of the political constraints they face and perhaps remain to be convinced how rigorous experimental Impact Evaluation can help them with those constraints. IE evidence is being used, but it is dependent on donor support, and at least half of the respondents see timeliness as critical.

2. IMPLEMENTING ORGANISATIONS

226. Just over 50 stakeholders responded to the section of the survey aimed at implementing organisations.

Evaluation funding

227. About a quarter of organisations have an independent evaluation department, but nearly half commission evaluation through individual programme teams. Separate evaluation commissioned by donors/programme funders is also common.²³ A badly worded question left it unclear how many organisations budget evaluation for their programme overall or by project. It would seem to be about half and half. Estimates of the appropriate budget allocation ranged from 3% of the programme/project budget up to 15% or even 25%. The majority fall in the range 3-10%. It was noted that USAID evaluation policy sets 3% as a minimum. 61% confirmed that they set evaluation budgets according to donor/ funder requirements. No respondents felt that their agency's evaluation budget was too much, and only just over a quarter thought it was 'about right.' However, the fact that the large majority felt it was 'a bit short' (46%) or 'not nearly enough' (28%) may reflect the fact that these stakeholders are evaluation specialists. Similarly, no one described their agency evaluation team as over-staffed, and over three quarters felt that staff numbers were not enough. Even staff capability was not rated that highly, with 45% rating it as weak or very weak.

Experience with Impact Evaluation

228. Implementer preferences on evaluation mirrored those of policymakers. Measurement of results and outcomes, monitoring delivery and reporting inputs were considered to be the most important M&E activities, in that order. Impact Evaluation was rated sharply lower, with experimental IE below non-experimental end of project evaluation. 12 of 23 respondents ranked 'Experimental (RCT)/Quasi Experimental Impact Evaluation' lowest, and only two ranked it highest. Respondents were also asked to rank which of their own Monitoring and Evaluation evidence had been most used by policymakers. Here the result was slightly different. Measuring results and outcomes remained at the head of the ranking but non-experimental end of programme evaluation was considered more important than monitoring from this point of view. However, experimental (RCT)/quasi-experimental IE remained with the lowest score.

229. Despite this, over 70% of organisations had programmes which included an experimental/ quasi-experimental impact evaluation. The most important reasons cited were to find ways of improving implementation, and checking if the programme is working. Making the case for funding and providing donors with evidence of what works were also mentioned by over a quarter. Requests from Government policymakers and researcher initiative were not important motivations.

230. Implementers rated the researchers who had carried out their IE highly. Only four of 37 felt they did a poor job. The reasons given were: 'randomization not being applicable', 'a lack of flexibility to context and lack of focus on utility'; and 'use of inappropriate goal and outcome indicators and inadequate measurement of outcomes'. These seem to indicate disagreement with the principle of IE more than with any performance issue.

²³ Multiple responses make these results slightly unclear.

Summary of implementing organisation responses:

231. Overall it seems that the implementing agencies share policymakers’ doubts about IE. Nevertheless, there are some who see them to be useful. The lack of budget and staff capability may be a constraint.

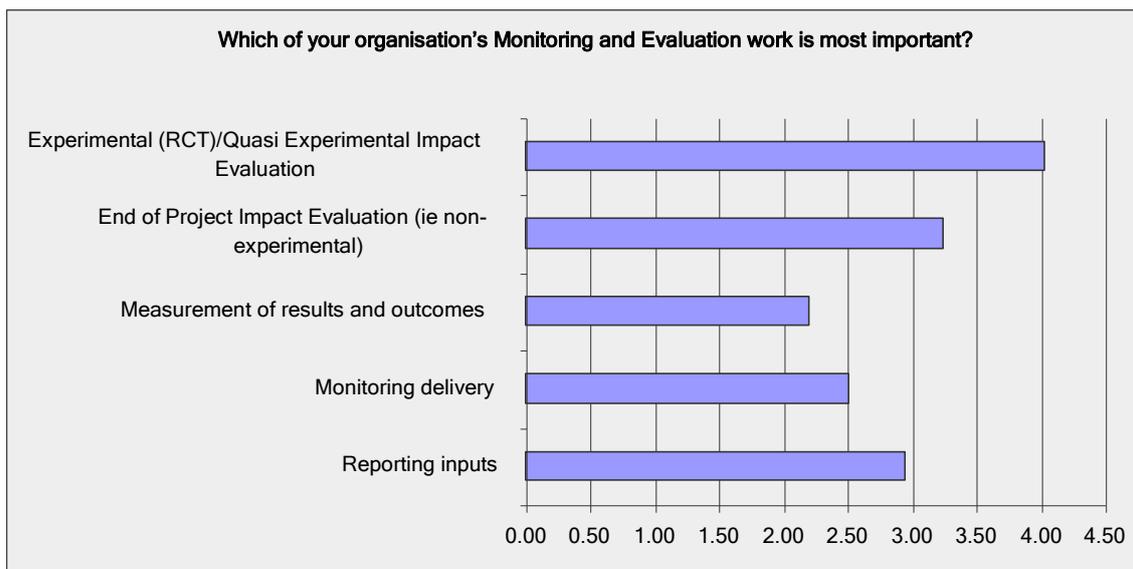
3. FUNDING ORGANISATIONS / DONORS

232. 18 stakeholders answered all questions in the section of the questionnaire aimed at donors and funding organisations. For many questions it was as high as 24.

233. Respondents reported that their organisations set aside less than 10% of their budgets for monitoring and evaluation, with a significant number budgeting 3% or less. The majority felt these budgets were ‘about right’ (46%) or ‘a bit short’. A quarter saw the budget as ‘not nearly enough’. For staff numbers, ‘about right’ changed places with ‘a bit short’ (50%). ‘Not nearly enough’ rose to 29%. Slightly more than half reported their agency staff’s capability to conduct evaluation as ‘adequate’ or ‘strong’, as opposed to ‘weak’/‘very weak’.

234. Donors were asked to rank the importance of the M&E work that they conduct. Their response was very similar to those of the implementers. Figure 9 shows the result.

Figure 9: Donor M&E Preferences



235. Once again measurement of results and outcomes, monitoring delivery and reporting inputs were rated as more important than impact evaluation, and once again experimental/quasi experimental impact evaluation was seen as being the least important by more than half of the respondents.

236. Asked to explain their reasons, 16 respondents replied. These written comments show how the emphasis on monitoring is partly driven by reporting requirements. Three comments help sum up the implementers’ views:

- *“Our work is about influencing government’ land reform policy through evidence of outcomes so measurement at that level speaks to the policy makers in no uncertain terms. This requires outcome measurement and quasi experimental impact evaluation.”*
- *“I believe the most value that can be added is through understanding outcomes. Inputs less important. Delivery should be locally accountable. Final impact is important but there is a long feedback loop. I have ethical concerns over RCT in its applicability in our programmes”.*
- *“Experimental design is not necessarily the best bang for buck in terms of improving our programming. - we can't wait 2-3 years to improve what we are doing, ... that is not good R&D, we have to use funds for process m&e, campaign m&e, questions of Program portfolio orchestration across multiple scales of work, organizational evaluations, in depth situational analysis of changing power actors, understanding cultural systems and institutions etc. Rigorous before and after studies actually address only a tiny percentage of an INGO's or bilateral aid organization's learning needs. ... the "steel rule" discovered in econometrics; the more rigor in the study, the more likely the study will show no impact, I have found this to be true in your reviews ... RCTs work best in projects in which the intervention is well controlled. Thus in a teleological way, you are simply suggesting that the best projects are the easiest to study. ... It ignores the idea that short term results very rarely predict long term deep change. This new trend in evaluation is doing very little to address the earlier debates about what makes change actually last when the funder leaves”.*

237. The last comment is quoted at length, with some editing, as being a good statement of the concerns 3ie will need to overcome in many parts of the donor and implementer communities, especially among INGOs. Rather than the resistance of naïve, conservative field workers, it will be necessary to persuade sophisticated and powerful thinkers with a well-supported point of view and a solid agenda to defend.
238. Despite these concerns, about two thirds of donor representatives said that experimental/quasi experimental evaluations form an important part of their evaluation programme. The most important reasons cited were ‘in order to better understand complex programme impacts’, ‘to confirm specific hypotheses about programme impacts’ and ‘to check whether programmes funded are working’.
239. About half of those responding had found it ‘more difficult than expected to find good evaluation teams’, or ‘a real lack of capable research institutions to conduct high quality IEs on development issues.’ The other half rated IE capabilities more highly.
240. 60 per cent of donor respondents come from agencies which fund organisations supporting impact evaluation. Apart from 3ie, these included J-Pal (5), IPA (7), NONIE (3), the World Bank (2), IFPRI (2), and Dev. Results (1).²⁴
241. Donor respondents were asked to describe their experience using IE evidence to support DC policy formation. 20% had not made the effort. Of those which had, over a third feel that DC policy makers welcome the evidence. Rather more feel that it has to be promoted and DC policy makers may see it as just one of many categories of evidence to be considered. In a comment, one noted, surely correctly, that it depends on the country.

Summary of donor responses

242. Overall donor representatives’ views mirror those of implementers and developing country policymakers. This can perhaps be summed up as saying that IE is important, but not as important as some other parts of monitoring and evaluation. For policy also, IE is part of a wider set of evidence to be considered.

²⁴ For unknown reasons one respondent said ‘prefer not to have them on your list’

4. EVALUATION PRACTITIONERS AND RESEARCHERS

243. With 281 stakeholders answering all the questions in the part of the survey aimed at evaluation practitioners and researchers (up to 320 for many questions), this group was much the largest in the survey. 75% reported that they had worked on programmes involving an experimental or quasi-experimental impact evaluation. As such, they represent a large body of knowledge and experience.

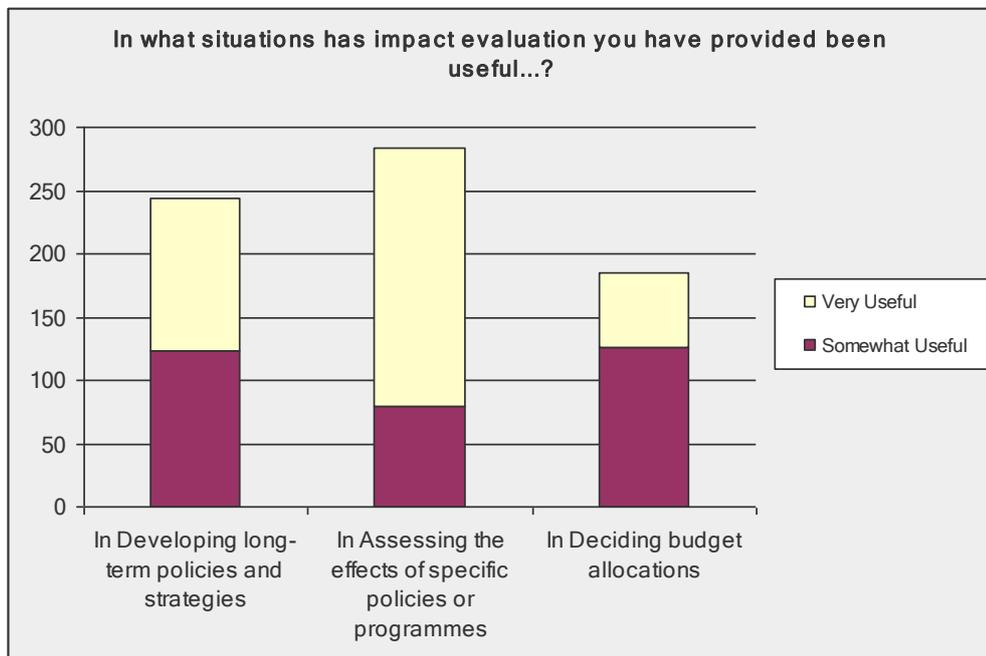
Experience with Impact Evaluation

244. Researchers and practitioners choose much the same reasons for IEs as other groups: ‘to check on whether the programme is working’, ‘to understand complex impacts better’, and to improve methods of implementation. Demand from Donors and Government was a factor but not so important.

245. When asked to rank which M&E work they used the most, researchers and practitioners gave similar answers as the other groups. As throughout, measurement of results and outcomes was rated highest. After that came non-experimental end of project evaluation and monitoring delivery. Although rated above reporting inputs, experimental (RCT)/experimental IE was still ranked as the least used overall. However, this result concealed a split between a group of 21% who rated it as most used and 40% who ranked it as least used. This split between proponents and advocates of experimental IE is reflected in later responses.

246. Overall, the low rating did not mean that most respondents felt experimental IE was not useful. Figure 10 shows how researchers/practitioners have found it useful in developing long-term policies and especially in assessing the impact of specific policies, but less so in budgeting.

Figure 10: Policy Use of IE Evidence – Researchers and Practitioners

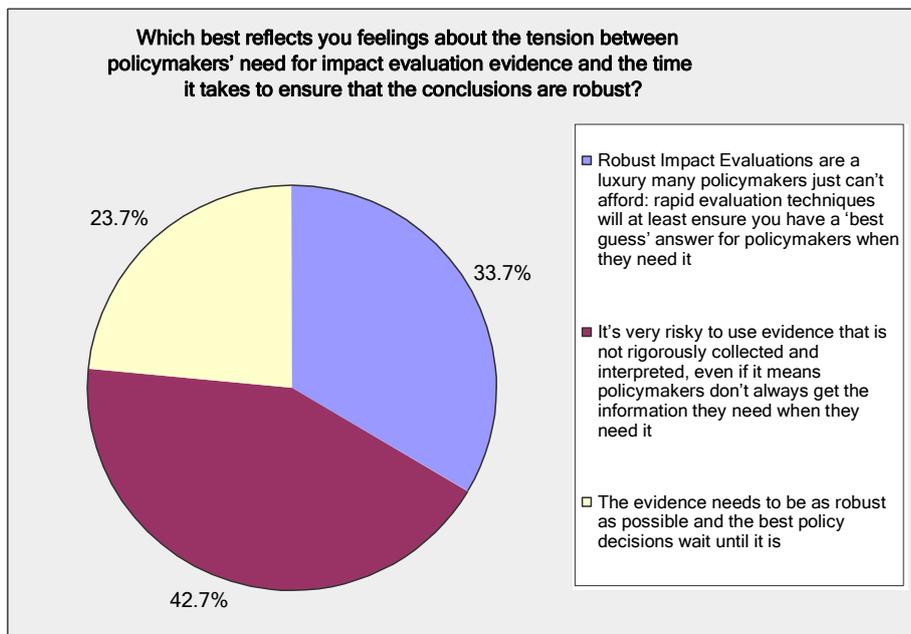


247. IE was also noted as being useful in eliciting debate, providing accountability, improving program design and implementation and testing potential solutions. It can serve as a basis for programme replication and scaling up, as well as for literature reviews and satisfying donor organisations.

Experimental Survey Design

248. Respondents were asked how much more difficult the complexity and context specific nature of work in developing countries makes it to implement effective experimental/quasi experimental designs. Over 60% responded that these designs are rarely implemented as planned, but that good management means useful results can be achieved. However, about a third thought that ‘a risk of failure or concealed biases mean there are relatively few situations in which it is practical to implement full experimental designs in developing countries.’ Budget constraints are also a factor.
249. On capacity, only 10% thought that ‘most countries have more than adequate capacity to use sophisticated statistical techniques’. The majority thought that some international support would be required, about a quarter suggest relatively intensive support.
250. Figure 11 shows responses to a question about the balance between the need for timely evidence and the time it takes to ensure the evidence is robust. About a third feel that it is better to deliver a ‘best guess’ answer for policymakers when they need it. However, the majority give priority to ensuring the evidence is robust.

Figure 11: Which is more important timeliness or quality? - Researchers/Practitioners



251. This question elicited 43 comments, more than any other. These reflected a number of themes:

- Many wished for an option combining best guess techniques when required with in- depth, robust analysis to follow up.
- The distinction between high quality evaluation and fast policy relevant analysis was seen by some as unhelpful. Or the underlying assumption that only experimental impact evaluations are robust as being false. A range of evidence is required, not just impact evaluations but holistic approaches and other tools. Or less forcefully, ‘It’s horses for courses.’
- A case of won’t rather than can’t for some NGOs, who think they know best. Time is a red-herring as policymakers seldom want to know evidence unless it supports their own perspectives. Other limitations include resources and budget, political context and timeframe. One commented ‘I would insert "won't" rather than "can't afford" and this includes

implementing organizations, specifically, NGOs, who are funded by donors on the basis of reputation and they are often not willing to introduce any sort of experimental design. In short, they are convinced they already know what works.’

- Robust IEs not considered a tool for policy makers, even though they find it interesting. One particularly harsh comment came that “3ie’s definition of robustness is political and ideological not scientific; policy makers are right to ignore 3ie.”
- The cost of full-scale IE as being a limiting factor.

Final Responses on Methods

252. All stakeholder groups were asked to answer two concluding questions on the mix of quantitative and qualitative techniques for impact evaluation. Figure 12 shows the response to the first question, on the balance between quantitative and qualitative methods. Figure 13 does the same for a question on the importance of RCTs.

Figure 12: Quantitative vs. Qualitative

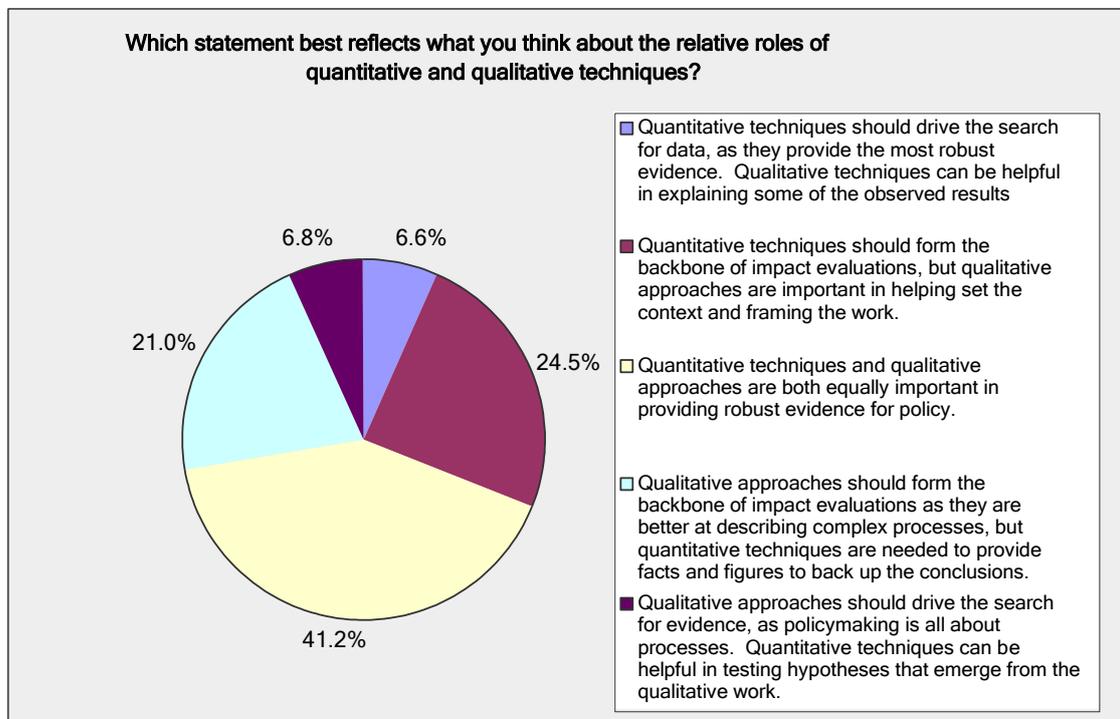
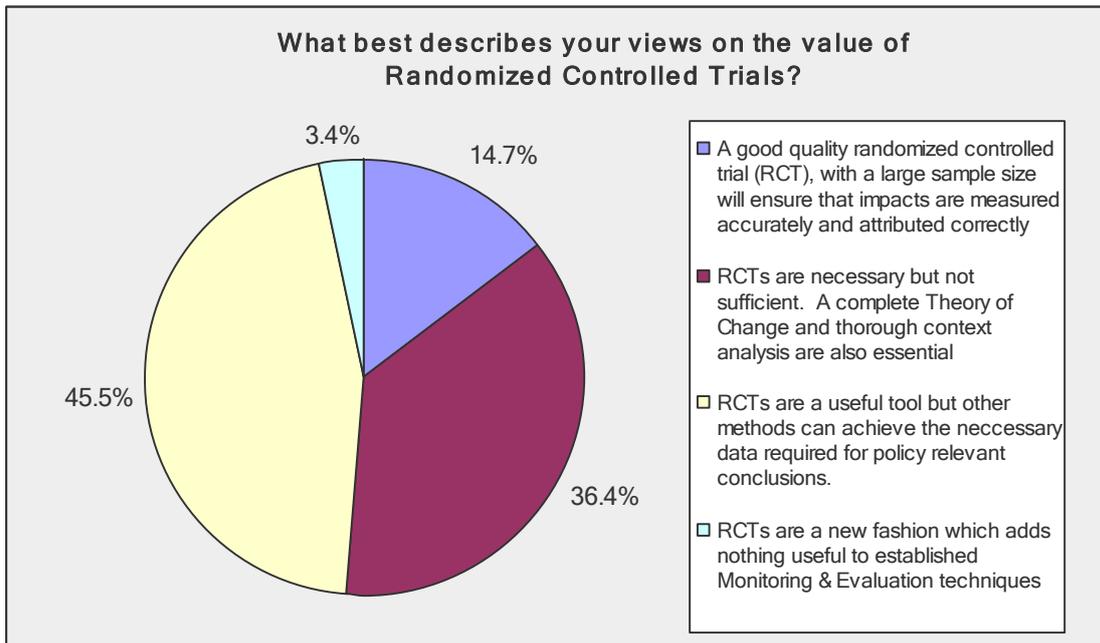


Figure 13: What are RCTs worth?



253. In summary, the two sets of results show a remarkably even split between advocates of quantitative and qualitative methods, with a large group in the middle who feel both are important. On RCTs, the split is even clearer, between those who see them as at least necessary, and those who feel they are at best a useful tool. Only a small minority see them as adding nothing useful.

254. This breadth of opinion was underlined by over 100 final comments on the survey. Some felt that 3ie’s perceived move towards qualitative methods (e.g. mixed methods and theory of change) was undesirable; whilst others felt it did not support these methods enough. Again there was a large middle group who feel it all depends; that different evaluation methods are needed depending on specific programmes and interventions, avoiding a one-size-fits-all approach. One respondent sums this up with the Einstein quote: "Everything that can be counted does not necessarily count; everything that counts cannot necessarily be counted". ‘RCTs are held as a golden standard, but often 2nd best techniques are needed to be used’ was another common statement.

255. Other comments were that there is a lack of support for less experienced researchers to access 3ie funding. The sharing of information and movement away from the mentality of a club of intellectual elites was another sub-theme to many stakeholder comments, as well as the lack of specificity to study conclusions.

4.4 Conclusions

256. The surveys of those working directly with 3ie provide a consistent picture. 3ie grantees are drawn from a group of experienced, well-funded institutions. In almost all respects, both Contract Negotiators and Principal Investigators are supportive of 3ie. If anything the relatively low level of complaints about the grant management process would indicate that 3ie is doing its job and providing the right level of challenge to keep grantees honest. For most grantees, however, it would

seem that 3ie is not essential, either as a source of funding or, in their view at least, as a technical resource.

257. The wider stakeholder survey was equally supportive of 3ie as an organisation. There was very high awareness of its work. Among its publications, SRs and Working Papers are highly rated. 80% use the website, and 50% find the IE and SR databases useful. Overall, more than half of respondents find 3ie's services and products useful in their own work, up to 60% for publications. The Expert Roster is the least appreciated service.
258. The division of the survey between different stakeholder groups was, in fact revealing of a broad consensus between policy makers, implementers, and funders on the role of Impact Evaluation. All groups give higher priority to measuring outcomes and monitoring delivery than they do to impact evaluation, and the more rigorous experimental approach to impact evaluation is the least important. With limited budgets and technical capacity, the expense and complexity of experimental IEs is a constraint. Although not expressed in these terms, these preferences may indicate that policy makers, implementers, even funders are not as concerned as 3ie's mandate assumes about the attribution challenge.
259. As the most important stakeholders of all, the views of developing country stakeholders are especially important. For them practical concerns and political constraints come first. Although the sample is small, it is a key conclusion that they remain to be convinced how rigorous experimental Impact Evaluation can help them with those constraints. They do use IE evidence, but it is dependent on donor support, and at least half of the respondents see timeliness as critical. Although 70% of implementers have had programmes subject to impact evaluation, their M&E preferences closely mirror those of the policy makers. They feel they lack capable staff for impact evaluation, and their evaluation budgets are largely determined by what their funders will support.
260. It is, perhaps, more unexpected that even donor representatives gave the lowest priority to rigorous, experimental impact evaluation. They too saw a lack of specialist staff as a constraint. Nevertheless, experimental/quasi experimental evaluation does form an important part of donor programmes: 'in order to better understand complex programme impacts', 'to confirm specific hypotheses about programme impacts', and 'to check whether programmes funded are working'.
261. Much the largest response to the survey came from researchers and evaluation practitioners. With over 200 having worked with experimental IE studies, they represented a large body of experienced IE practitioner opinion. Yet there were clear divisions on a number of points. Overall, like the other groups, they rated experimental IE as the M&E technique they used least. However, this concealed a split between about a fifth who rated it most used and two fifths who rated it as least used. This split was repeated in later responses, more or less 50:50 or 60:30:
- A majority felt that experimental IE is practical in developing countries, but about a third chose the response: 'a risk of failure or concealed biases mean there are relatively few situations in which it is practical to implement full experimental designs.'
 - On the balance between policymakers' need for timely evidence and the time it takes to ensure the evidence is robust, once again a third chose the response: 'Robust Impact Evaluations are a luxury that policy makers cannot afford' (This question elicited 43 comments, many strongly worded, to underline the strength of the division amongst evaluation practitioners.)
262. Two final questions sought the views of all stakeholder groups on two fundamental technical questions, the balance between qualitative and quantitative methods, and the value of RCTs. Although the responses covered the whole spectrum, there was a clear consensus that a mix of methods is needed. The consensus was nicely balanced between those who give primacy to quantitative methods and see RCTs as 'essential', and those who give primacy to qualitative and see RCTs as 'useful'.

263. The aim of this survey was not to enter into the rather sterile debates about ‘randomistas’ and the ‘results oriented agenda’. Instead the purpose was to see how far 3ie’s own message about Theory-Based Evaluation and mixed methods has won support, and crucially where it should now focus the technical direction of its work. As an organisation, it has won wide acceptance and created a strong platform as an authoritative voice on impact evaluation. From the comments on the survey, there are still many who see 3ie as an advocate for rigorous quantitative trials. To the extent that 3ie focusses on rigorous evidence to address the attribution challenge, this is correct. However, it means that the other equally important part of 3ie’s message, about theory-based evaluation, has not been so widely received. The evaluation team believes this is reflected in the fact that policymakers, implementers and donors all put experimental IE at the bottom of their list of priorities: implicitly rating it as ‘useful’ rather than ‘essential’; and in the fact that 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 more than just useful, and that Theory-Based Evaluation is the way to make it so.
264. Returning to the grantee part of the survey, the one clear exception to their general appreciation of 3ie’s work concerns the Policy Influence Plans. Respondents were divided between a group who are comfortable with policy influence activities, even feeling that there is not much 3ie can teach them, and those who are distinctly uncomfortable with the whole approach. The need for frequent letters of stakeholder support was a particular issue. Some see them as actively counterproductive. Together with the survey information on what types of evaluation are seen as important for policymaking, these results do suggest that 3ie needs to carefully consider whether asking all its projects to demonstrate policy impact is an appropriate strategy.

5 Process Evaluation

265. This chapter develops the status report presented in Chapter 2 in order to assess the economy and efficiency of 3ie’s processes. The assessment aims to be forward looking, concentrating on the processes as they currently stand, rather than on past issues. Six processes or groups of processes are considered:

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

266. To aid an assessment of the economy of 3ie’s processes, section 5.1 presents an analysis of how 2011 expenditures were allocated across different activities and cost categories.

5.1 3ie Expenditure by Activity and Category

267. Table 19 presents an analysis of 3ie’s expenditure for 2011, broken down by major component and category of expenditure. Some adjustments have been made to the original data so as to provide as accurate a picture as possible of the way 3ie allocates its resources. Every year is different, but data is not available to make an analysis over the longer term. A comparison with the audited figures and projections shown in Table 14 above indicates that 2011 was reasonably typical. With one significant exception, it matches the budget for 2012 quite closely.

268. There is one possible weakness in this analysis: the possibility that some activities generate revenues to offset the expenditure. Expenditure on sponsored conferences is an important example, but this is recorded separately. In the table it is treated as being equivalent to a grant. It is known that there are other revenue generating activities, but it has not been possible to identify them, or the offsetting revenue. In 2011 at least, they are not believed to have been large. This may change as 3ie starts to operate large dedicated programmes, such as the \$20+ million BMGF HIV/AIDS grant. These include an element to cover 3ie’s management cost arising from with those programmes.

Table 19: 3ie Expenditure 2011 by Component & Category

	Component					Total	%
	1	2	3	4	5		
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	

Notes 1) \$3 million paid to GDN but not disbursed in 2010, added to Cpt 1: Grants

2) GDN fee allocated across components 1, 3, and 4: 80:10:10

269. Apart from grants, the table shows that 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.
270. The last line in the table analyses the distribution of non-grant costs across the components. As the best indicator of 3ie’s current strategic balance it is presented again in Table 20 compared with the 2012 budget and the distribution proposed in 3ie’s Strategy 2010 – 2013.

Table 20: Relative Weighting of 3ie's Components

	Non-Grant Cost 2011 %	Total Cost 2011 %	Budget 2012 %	Strategy 2010-13 %
1. Evidence Generation	30.7	75.4	65.7	77.8
2. Evidence Synthesis/Dissemination	12.0	4.3	4.1	3.9
3. Building an Evidence Culture	17.6	6.8	8.4	1.6
4. Developing Evidence Capacity	9.3	4.5	3.8	1.2
5. Institutional Capacity	30.5	9.0	18.1	15.6

271. The table shows significant shifts in the strategic allocation: a substantial and continuing increase in the allocation to Component 3 and, to a lesser extent, Component 4. As it is effectively an overhead cost, the low figure for Component 5 in 2011 seems positive, while the sharp increase in the 2012 budget seems negative. However, the jump is partly by a substantial new activity supporting and expanding the membership and funding base; and, the budget for the impact and process evaluations of 3ie itself.
272. The proportion of expenditure which does not go to grants is a standard measure of the economy or efficiency of a grant-making organisation; although it must be used with care. Every organisation provides a different combination of grant-making and other services. In 3ie’s case, for example, Components 3 and 4 are clearly non-grant making. The economy of their operations must be judged on a different basis. To make the calculation properly comparable, the overhead costs of Component 5 have been allocated pro rata across the four active components. Table 21 summarises the results.

Table 21: Analysis of Overhead Cost Distribution

	Component			
	1	2	3	4
Total Expenditure	10,317,363	585,339	936,175	616,311
Cpt 5 Distributed	543,599	212,104	310,815	164,986
Total loaded	10,860,961	797,444	1,246,990	781,297
% Total Allocation	79.35	5.83	9.11	5.71
Non-grant loaded	1,782,113	695,355	1,018,963	540,885
% Non-grant	44.14	17.22	25.24	13.40
	Non grant ratio - %			
Total	16.41	87.20		
Ex GDN/Cpt 5	5.01	60.60		
	543,620	483,251		

273. 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. The 87% non-grant figure for Component 2 should not be taken to mean that the Component is not being managed economically. Instead, it highlights the need to define clearly what each component's strategic role should be. For Component 2 in its current form, the justification must be as a standalone activity, not through grant making. It is also possible that the ratio will change sharply when the larger SR3 and SR4 programmes start to operate.
274. Component 1, on the other hand, clearly stands or falls by its efficiency as a grant-making operation. Especially so since a number of activities which add value to the basic grants – dissemination etc. – are covered under other components. On that basis, 16% must be considered high. However, over 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 Component 1 comes down to just 5%.

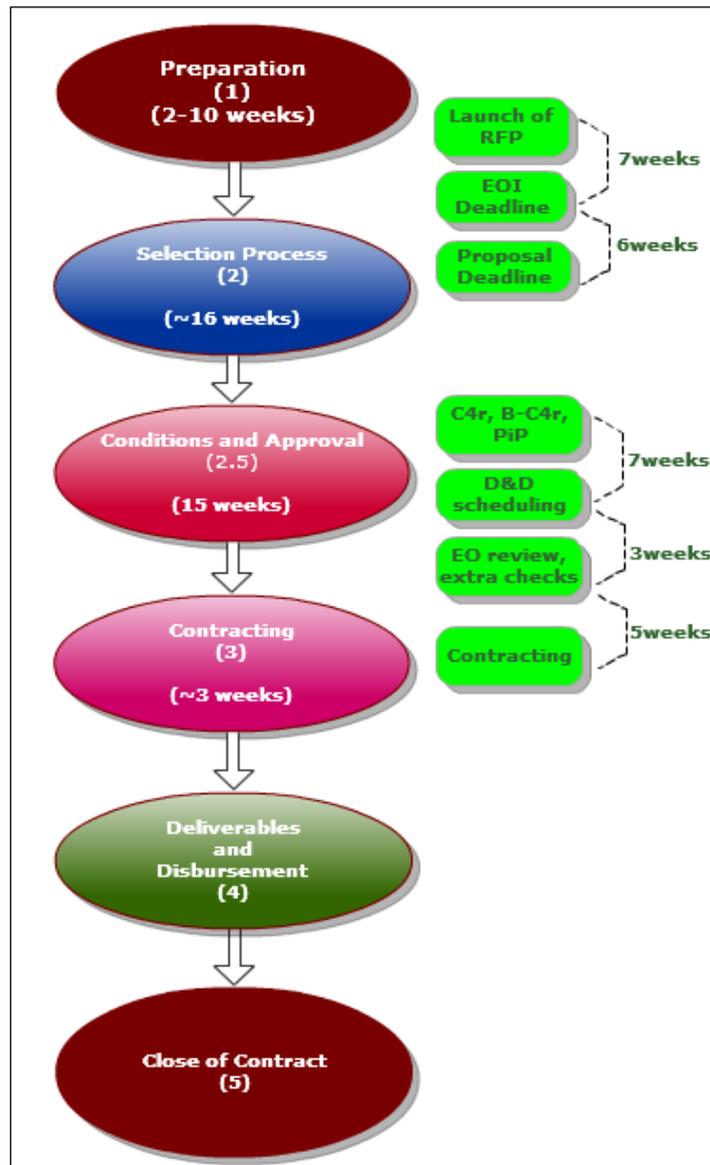
5.2 The Grant Management Process

275. 3ie's grant management process is coordinated by the Programme Office (PO). The PO manages all communication with the grantees and with the contracted technical reviewers; and it coordinates the actions of the Evaluation Office, PACO and GDN's finance unit. Initially it was expected that GDN's programme office would have this role, and they managed the early phases of OW1 and SR1. In 2009, the GDN staff member who was managing the portfolio was transferred to 3ie. Since then 3ie's own unit has grown to have four staff. GDN's role as the formal contract holder is now limited to the approval of budgets and managing grant payments. All other actions are driven by the 3ie PO. 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.

5.2.1 3ie Grant Procedures

276. In June 2011 a Grant Procedures Manual was issued. This sets out the current process, which is summarised in Figure 14. This process has been developed from 3ie's experience with Open Windows 1 to 3.

Figure 14: The Grant Management Process



277. The selection of grant applications is the most critical and intensive part of the process. It is reviewed separately in Section 5.2.2 below. After selection, the PO coordinates a process of negotiation and adjustment (Conditions and Approval in the figure). Applicants are asked to consider a set of Comments for Response (C4r) and Budget Comments for Response (B-C4r). The C4r is prepared by the EO and the EPA nominated for the grant, and the B-C4r by the GDN finance office, together with the 3ie Finance Officer. PACO reviews Policy Influence Plans (PiP) at this stage as well. Subject to satisfactory responses to the issues raised in the C4rs, the final grant agreement/contract is drafted and sent to the applicant for signature.

278. Each agreement includes a full schedule of disbursements (tranches) and deliverables, against which those tranches will be paid. As well as regular progress reports, grantees are expected to submit specific technical deliverables: draft questionnaires, survey plans, datasets, interim analyses etc. The EO decides what deliverables are appropriate. When they are submitted, the EO and the EPA check that the deliverables have met the requirement before approving the tranche payment, which is also cleared by the 3ie and GDN Finance Office staff.

279. Figure 15 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 latter is measured by when a significant proportion of grants has been completed, or is due to be complete. The figure shows that from launch of the major open windows it took five or six months 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. In total, it took between 11 months (OW2) and 14 months (OW3) to complete the whole process. For SPTW, time between grant agreements and first tranche payments has been 11 months or less, an improvement over OW3.
280. 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. OW3 was the slowest of all. It is possible this was because of the introduction of an additional EoI step in the process and the migration to the on-line GMS. However, in looking at these timelines it is important to take the scale of the task into account. This should become clear in the next section, which describes the selection process.

5.2.2 The Selection Process

281. From 79 applications for 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 thorough technical review of this many proposals is a major exercise, especially so 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 ESSRC.
282. For Impact Evaluation windows full proposals are scored by at least four reviewers, usually one internal and three external. These scores are then used to select a long list of grants to go to the Review Panel. The scale of the task is measured by the fact that 94 grants went to the OW2 review panel, and 62 to the OW3 panel. For OW2 it took a panel of 17, plus four 3ie staff, two days to complete the selection: starting by a sort into 'Yes' and 'No piles' followed by voting for or against those in the Yes pile.
283. The same basic scoring framework has been used throughout:
1. Technical quality
 2. Experience of nominated staff
 3. Involvement/capacity building of developing country researchers
 4. Involvement of 3ie member/associate
 5. Policy relevance
 6. Potential policy impact (OW2 onwards)
284. For each criterion, scores of between one and ten indicate whether they are 'not met' (0 – 2); 'weakly met' (3 – 5); 'sufficiently met' (6 – 8); or 'exceeded' (9 – 10). A score over 7 is considered technically qualified for submission to the Review Panel. Applications scoring between 6 and 7 are reviewed again to decide whether or not the Panel should see them.
285. 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 OW1 policy relevance only had a weight of 10%, compared with 40% for technical quality and 30% for staff experience. For OW2, the addition of a Policy Impact score raised the total weight on policy to 30%. For OW3, a separate review by policy specialists was added. They looked at all technically acceptable proposals: 61 out of 204. The fact that the technical scoring already included the technical reviewers' mark for policy relevance/policy impact element implied an even greater weighting towards policy.
286. However, there are signs that the technical review results were still the most influential. In one case, 'recommend rejects' from both the policy reviewers were over-ridden and a grant awarded. It is indicative of the challenge of selecting proposals which maximise both policy relevance and technical quality that a third approach was introduced for SPTW. The separate policy review was abandoned but the reviewers' technical and policy scores were weighted and averaged separately to produce two scores, one technical and one policy, for each proposal. It is perhaps not surprising, given that both scores were given by the same reviewer, that there was more consistency in the results this time.
287. To illustrate the overall methodology, the weightings used for SPTW are shown in Table 22. This was the first review to include a score for cost, somewhat awkwardly summed into the policy score; and the first to add a distinction between the internal and external validity of the technical proposal. Internal validity covers details of the study design, including gender awareness, while external validity considers the *"incorporation of the intervention's theory of change into the evaluation design, the use of mixed-methods as appropriate, and cost-*

effectiveness or cost-benefit analysis. Reviewers should also assess the way in which the design ensures policy relevance and research uptake.”

Table 22: Score Weighting for Grant Proposal Reviews - SPTW

Criterion	Weight
Technical Score	
1. Research Team Qualifications	20
2. Involvement of Developing Country Researchers	15
3. Internal Validity	30
4. External Validity	30
5. 3ie Membership	5
Policy Score	100
5. Policy Relevance	40
6. Policy Impact	40
7. Cost	20
	100

288. With some variations, the selection of SRs is similar. For SR1 and SR2, ‘adherence to Campbell Collaboration procedures’ and ‘incorporation of realist perspective’ were specific requirements, with a 50% weighting between them. These have been replaced with greater weight on relevance and dissemination and developing country involvement.
289. For external reviewers and external peer advisers, 3ie draws on a panel of specialists. The master list includes 139 specialists, some of them very eminent indeed. US universities which have made experimental IE a focus of their work are strongly represented: Harvard, Berkeley, MIT and Yale. However, other institutions with several names on the list include IDS, Oxford University, IFPRI and the World Bank. The reviewers are identified as specialists in different technical sectors: health, agriculture, education, etc.
290. In practice, most of the reviewing has been done by a smaller group, not all of them even on the panel list. Up to Q2 2012, 3ie used 108 different reviewers, but 54% of the work was done by a core group of 20. These were drawn from a wide range in institutions, less than half from US universities and only a minority from the major institutions just mentioned. The reviewer who did most work of all was from the University of Cape Town, and UK and European universities were well represented.
291. 3ie has a formal Conflict of Interest policy, which applies to reviewers. This does not prohibit conflicts, but it requires them to be managed appropriately. In that context, some 16% of PIs on 3ie grants have acted as reviewers at some point, as have four past/present Board Commissioners, at least one of whom has also acted as a PI. IE studies require highly specialised skills and it is not surprising, or a point of particular concern, that the pool of qualified reviewers available to 3ie is the same as the pool of capable grantees. In practice, 3ie has thrown the reviewer net very wide. The one exception perhaps concerns SPTW, where a relatively narrow group of grantees does overlap with the panel of external peer advisers. IFPRI has four of ten grants under this window, and IFPRI staff members are named as External Peer Advisers on the grants to other organisations. One is acting as Lead PI on one grant and EPA on another.

5.2.3 The Grant Management System

292. Apart from publicity on the website, OW1 and OW2 were managed using email and spread sheets. In April 2010, 3ie decided acquire an on-line Grant Management System (GMS). The ToR set out a very ambitious specification. It asked for every stage of the grant management process to be brought onto an integrated online GMS, starting with the receipt of applications and ending with grant disbursement, M&E and reporting. Requirements included such specific details as the generation of letters of regret to unsuccessful applicants, the coordination of reviewers’ comments on the applications, and direct on-line access to applications for Review Panel members. In summary, every aspect of the management process was to be managed on-line.
293. Of three bids received, the contract was awarded to Praxis, a South African company which was offering its web-based Project Tracking System. The bidder claimed that this already met most of 3ie’s operational requirements. The following table is a summary of what has happened since:

Table 23: Summary of GMS progress

March 2010	Praxis bid selected. Offer price for 3 Phases -	\$95,639
April 2010	Effective date of contract	
	Negotiated price to end of Phase 2 –	\$165,553
	Main delivery date -	August 2010
June 2010	Contract signed by 3ie	
June 2010	LoV1 signed	
	Price correction, added application pack for OW3	
	Revised price -	\$228,053
September 2011	LoV2 signed	
	New ‘Workflow Module’ etc., PW+TW packs	
	Contract extended to 30/9/12	
	Main delivery date -	Not specified
	New price -	\$431,275

294. The Evaluation Team have reviewed the GMS documentation briefly. Although they have some experience managing IT projects, they are not IT specialists. Subject to those caveats, they would register significant concerns about the GMS, starting with the substantial drift in both cost and timeline. What can be described as the front end of the system is operational - registering and managing grant applications and the application review process. Although not all applicants are happy with it, 3ie staff report it is working well. If, however, there are any changes at all in the way 3ie wants to launch a new window, Praxis has to write a special application pack, as noted in the list above, at a cost of between \$10,000 and \$30,000 per window, and potential delay to the launch.
295. The back end process, i.e. the management of grants after award, has still not migrated to the GMS. The contractor has advised that this will require the purchase of a new software package and its integration into the GMS, to create the ‘Workflow Module’ which forms the major component under LoV2. At the time of the evaluation team’s visit to Delhi, it seemed far from certain that this would be ready by the contract end date in September 2012.

296. A number of factors have contributed to this drift:

- The clear and detailed specifications set out in the 3ie ToR were not carried through as solid deliverables under the Praxis contract. The Praxis proposal itself was a rather general response to those ToR.
- 3ie has fallen into the common trap of asking for additional functionalities before the core specification has been met.
- The original requirement was extremely ambitious. Managing IT projects at this level of complexity is a specialised skill.
- Managing the IT contractor forcefully is a key part of that skill. 3ie does not seem to have done this.

297. More generally, it can be noted that IT systems have the greatest potential when dealing with high volumes of repetitive transactions. This is not the case with 3ie’s grant management function, something highlighted by the need to re-write the GMS for each grant window.

5.2.4 Comparison with Other Grant Making Bodies

298. The Evaluation Team are asked “to carry out a comparative study with similar grant making bodies, for comparison in terms of review processes, policy engagement and so on”. There has not been time to visit and interview an equivalent body. However, the Team Leader has himself managed a £3 million a year research grant programme for DFID and reviewed a number of civil society grant funds as well. The grantee surveys provide additional evidence on this point.

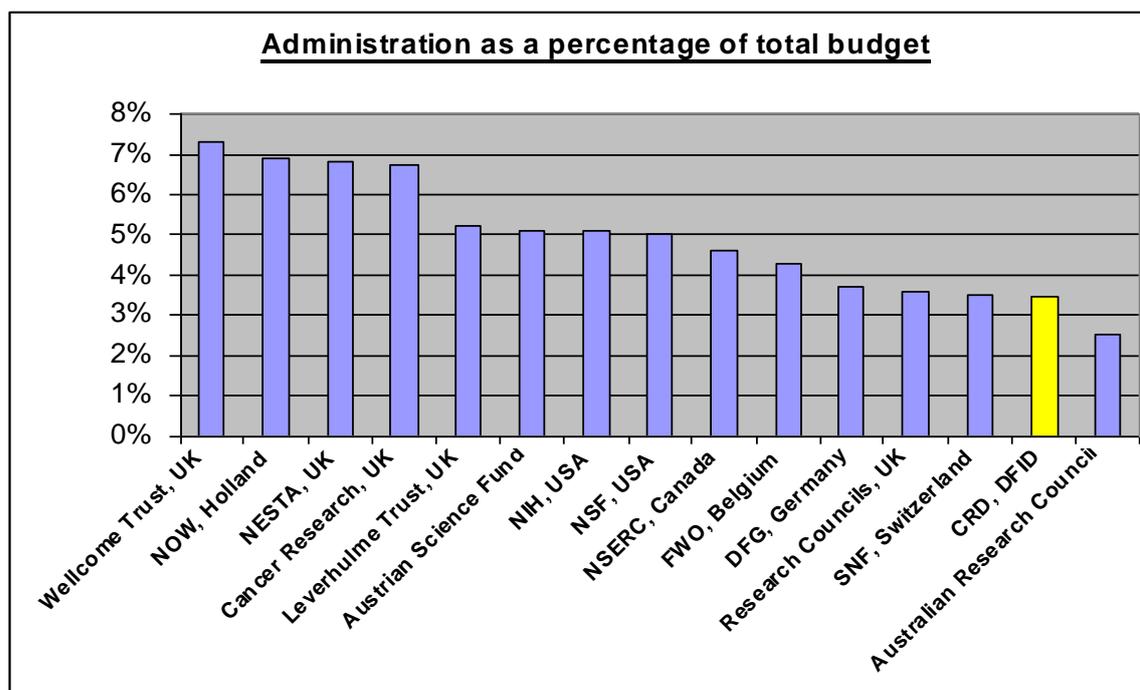
299. With regard to review processes, 3ie has worked to an unusually high standard: in its selection processes; and, especially, in its management of Deliverables and Disbursements. There are few other organisations which monitor progress and the quality of the work so closely. 3ie also gives unusually high emphasis to accountability, sending rejected applications for independent review, even at the EoI stage. While not on the scale of much larger organisations such as the UK’s Economic and Social Research Council, and the USA National Science Foundation, 3ie’s guidance to grant applicants is comprehensive, and at least as accessible as that provided by the National Science Foundation.

300. Policy engagement is discussed at length elsewhere in this report. Suffice to say that in this respect 3ie faces a much bigger challenge than research funding organisations working in the developed world. Only funds supporting civil society advocacy in developing countries face a bigger one. Although they are not grant funds, DFID’s Research Programme Consortia are perhaps the best benchmark for 3ie in this respect; described in DFID’s Terms of Reference as Communications: Getting Research into Use. A quick review of one RPC’s website, EdQual does not show a focus on policy influence equal to 3ie’s efforts under Component 1 and Component 3. Instead the emphasis remains on communicating research results.

301. It is not easy to benchmark grant management costs to determine what is reasonable. For large civil society funds disbursing large sums at infrequent intervals to grantees who are expected to deliver without close support or monitoring, it is possible to reduce the management overhead to below 5% of the total spend. For a fund, managing large numbers of small grants, to grantees who require capacity building and steady support, the grant management cost rises sharply. Setting the boundary between purely administrative costs, and value-added services such as advice, dissemination, etc., is always a somewhat arbitrary decision.

302. Figure 16 shows some comparative costs prepared by DFID. These can be compared with the figures for 3ie shown in section 5.1: 5% for the grant management function, without Component 5 and GDN overhead, 16% with that overhead. However, this comparison needs

care. The organisations shown in the figure are all very large funders of first-world research. In other words they are dealing at large scale with long-established organisations, working in a long-established framework. To take the UK’s ESRC as one example, it has a £250 million budget, much of which is spent in multi-million pound grants to established centres. Barely a 10th of its budget goes to grants between £200,000 and £2 million, similar to those awarded by 3ie. Any firm conclusion on performance would need a detailed comparison of how ESRC and 3ie deal with this category of grant.



Assessment

- 303. 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 are 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 the technical assessment to include the quality of the intervention theory of change and the use of mixed methods. Cost is now also part of the scoring. How far these changes have improved the quality of the work 3ie supports is reviewed later, in Section 6.3
- 304. At around a year from RfP to signature of the grant agreement, processes are slow. 3ie is managing very large numbers of grant applications and it is demanding high standards of technical quality and financial control of its grantees; higher than most other grant-giving organisations. That said, the evaluation team believes there is a need to ensure that the timeline set out in 3ie’s Grant Procedures Manual is met. However, for reasons discussed elsewhere, the Evaluation Team would emphasise that this should not come at the expense of the standards being applied.
- 305. There are two ways 3ie might consider streamlining the process. One would be to make a harder cut at the EoI stage. At present 3ie only rejects EoIs which are clearly non-compliant or under-qualified. A rule that only, say, three times as many EoIs will be accepted as 3ie has

funds available for would save both 3ie and many grantees wasted time. 3ie feels that even rejected EoIs should be externally reviewed. The Evaluation Team feels this may be taking fairness too far. The second change would be to require a fuller final proposal, requiring applicants set out their ToC and their context analysis in more detail. As discussed in Chapter 6.1, the Evaluation Team believes there is a clear need for this.

306. The Evaluation Team are concerned that the effort to develop an integrated on-line Grant Management System is proving to be an unnecessary distraction for a critical function which is working well in its current, relatively low-tech form. It is worth pointing out that at the current contract value, 3ie could have doubled the staff of the Programme Office for 10 years with the money being spent on the GMS. It is recommended that the business case for the GMS be reviewed on a zero-budget basis. One option might be to discontinue the back-end, Workflow Module, and focus on completing the application/grant selection front-end. This is the part of the process where volumes are larger and most tasks are relatively repetitive. Eliminating the need for an expensive re-write for every new grant window should be a priority.

5.3 Policy and Advocacy Processes

307. The PACO team has been an integral part of 3ie from the outset. It is mainly (but not wholly) responsible for components 3 & 4, contributes to component 5, and has also played a significant part in the grant processes with the design of the Policy Influence Plans (discussed in section 3.3.2). From a single individual, who joined in early 2009 and covered all aspects of the work, the team now consists of four posts, split between the policy advocacy and the communications aspects of the work. PACO's work has been informed by the Learning and Advocacy Strategy published in early 2010, and the Knowledge Management Strategy published in October 2011.

5.3.1 External communications and knowledge management

308. External communications involve the dissemination of information to outside audiences, whether those be developing country policymakers, the evaluation community or development practitioners more widely.

a) The website and e-newsletter

309. These are considered together as they are 3ie's main platforms for on-going dissemination. The website was originally developed in Cairo at a cost of \$4,500 as a bespoke platform. Management was handed over to an Indian firm for a monthly retainer of \$500 in 2010. In 2012 it was revamped in line with recommendations from the KM strategy. As well as being the main repository of electronic versions of 3ie publications, the site offers information on 3ie projects and grant windows, other funding opportunities for evaluation, training opportunities, evaluation-related events and job announcements. It also publicises the expert roster and systematic review database. It thus addresses 3ie's desire to be a knowledge broker for IE-related evidence, not just a simple clearing house of information. In the past, most users reported visiting the website for grants information. In line with recommendations in the KM strategy, the redesigned website attempts to guide users to specific content targeted to their likely need.
310. The website is run by a team of two who focus on editorial work; ensuring that there is a good supply of copy both from 3ie's projects and from the wider IE community. This is a complex process as it comprises both reporting and magazine-type editing. Revamping the website occupied a good deal of their time in early 2012: now that it is in place, more will be done to pull stories out of 3ie-funded grants as individual projects come to an end. The

biggest resource for the website is the IE and SR databases; the Communications team writes features for the website analysing the findings from different IEs and SRs (not all conducted by 3ie) and also invite other evaluators to blog or comment.

311. The e-newsletter reflects the content of the website, being a mix of reporting and analysis/magazine-type editing and driving people to the website for further information. It is currently sent out to approximately 8,000 people via a mass dissemination service – approximately one quarter of those are thought to open and read it, though this cannot be independently verified. The plan was for a newsletter every two months, though it is now being sent out quarterly instead. In future the plan is to increase the amount of analysis-based writing in the newsletter. The KM strategy recommended various changes to the newsletter to improve its ability to function as an online marketing tool, though this does not appear to have been taken up as yet.

b) Publications

312. The various publications produced by 3ie are outlined in section 2.4 above. The Communications team supports the publications process; editing and proofing all the Working Papers, Systematic Reviews and Policy Briefs. The editing workload depends on the ability of others to produce drafts for comment; these may include SRs (from the London office), project reports (from grantees) and WPs (from the Programme Office). It has therefore been difficult to develop a strategic approach to communicating individual publications though this may change in future as more project outputs are finalised and fed into the Policy Briefs series. Detailed data on processes and costs for the major publications series are not available.

c) Journal of Development Effectiveness

313. 3ie and Taylor Francis have published this quarterly since 2009: the ED is the Managing Editor who works with an editorial board of nine people. From the outset this journal has been extraordinarily successful, as evidenced by the recent granting of an ISI factor of 0.92 – a very high score for such a young journal. The initial proposal was very well received by publishers; not only did it fill a clear niche for a journal related to evaluation, but it was seen to be forward-thinking in two areas – reporting negative or inconclusive results, and addressing the issue of policy impact. Taylor Francis were so confident of its impact that the journal immediately appeared quarterly, rather than the usual 3 times per year for a start-up. There has been a very natural flow of submissions, based in part on 3ie's extensive network and effective use of e-media. Download statistics for 2011/2 are not yet available, but the KM strategy reported that it was downloaded approximately 4,000 times in 2010. The journal has always been published to page and budget and has a good spread of authorship from across the globe. 3ie's early communications activities, such as establishing networks, the newsletter and the website were all instrumental in ensuring that the journal had a wide reach; 3ie pulled the journal into on-going conversations with the IE community. From Taylor Francis' point of view, 3ie were one of the few organisations who recognised the importance of developing a co-ordinated communications strategy involving multiple interlinked platforms.

d) Seminars and conferences

314. The two biennial conferences and the seminar series in Delhi, London and Washington are major aspects of component 4 – building a culture of evidence. Responses to the post-Cairo survey were broadly positive and, as noted earlier, the survey sent to the Cuernavaca

participants 3 months after the conference showed that participants appreciated the conference and its mix of topics covered. Seminars are put on locally by the Delhi, London and Washington offices: the PACO team organises the Delhi conferences but has little involvement in organising and running the seminars in London and Washington. The topics for seminars are difficult to decide in advance as it does depend on who is available or happens to be passing through. It is also not possible to estimate the proportion of staff time spent on organising the seminars without detailed data.

315. The evaluation team analysed the feedback forms from 14 seminars held in Delhi between November 2009 and July 2012, and 4 held in London held in 2011. The number of forms filled in varied from 2 to 17, with an average of 6. It was not possible to correlate this with the number of attendees to get an indication of the overall response rate. Feedback on the whole was positive with most attendees agreeing or strongly agreeing that the seminar was clear, useful, and would inform their work on impact evaluation. Negative comments were mainly about the degree to which participants were interactive during the seminar, with a few saying they found it difficult to fully comprehend the subject matter, which varied between seminars. There was a wide range of suggestions for future topics, indicating participants' on-going interest in the topic.

5.3.2 Demand generation workshops

316. Over 20 demand generation workshops have been held. Most have been held in developing countries though a couple were held in Europe for DFID and SIDA. The ED carries the bulk of the presenting load, accompanied by local IE expert practitioners. The intention is to familiarise policymakers and practitioners with the need for IEs; to encourage institutional affiliation with 3ie and ultimately application to 3ie's Policy Window. The workshops have overall been well received; participants scored them consistently above 75% for relevance, clarity and the level at which they were pitched. Interviews with IPA confirm that these types of workshop and seminar do generate a great deal of enthusiasm for IE approaches, though for institutional and political reasons it is often hard to carry this through to implementation. 3ie does not explicitly report the payoff from demand generation workshops, either in terms of the numbers of PW applications generated, membership applications or other impacts, so it is not possible for the Evaluation Team to judge how effective these are in building the culture of evidence over the long term. Section 7.2 notes that expanding the remit of these workshops to focus more on advocacy and service provision than on generating fundable IEs may give greater returns in terms of generating wider membership.

5.3.3 Internal communications

317. Internal communications cover the Board of Commissioners, Members, Associate Members and partners. The 3ie workflow document (annexed to the Learning and Communication Strategy) shows that PACO has limited responsibility for internal communications to these groups; apart from the newsletter (already mentioned) and preparing the annual report. The ED, DD and PO have most responsibility for communication with the Board. The External Liaison Officer is responsible for communicating with 3ie's Members and Associate Members, though as noted in section 3.2 (membership) the job description is mainly about ensuring compliance with institutional issues.

5.3.4 Communications with grantees and the PIP process

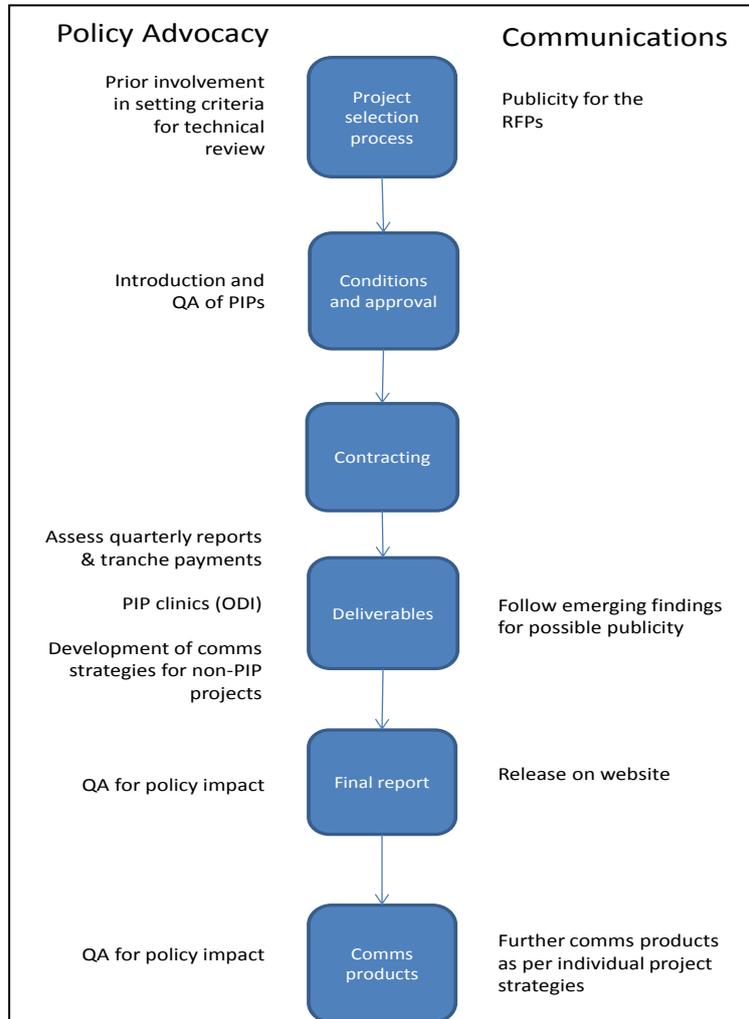
318. The diagram below summarises the various activities PACO undertakes to communicate with grantees. In the initial stages of a grant window, the team supports the Programme Office to publicise RFPs widely using its extensive list of communications channels. It has built up a useful targeted list of 17 listservs, 40 portals, 22 evaluation networks, 8 development

economics networks and 24 other communications channels. Newsletter recipients are also re-mailed when a call is being launched. The combination of both mailing lists gives 3ie a very wide reach to potential grantees.

319. In spite of the emphasis put on policy relevance in the scoring system for selection, PACO is not involved in the selection process (the number of proposals received would swamp the small team) though it commented on the instructions given to reviewers as they were being drafted. PACO involvement with grantees begins after selection has taken place, during the Conditions and Approval process. Successful grantees are asked to prepare a Policy Impact Plan (PIP) to a specified format. As noted in section 2, PIPs were introduced during OW2. During OW1 and the beginning of OW2, all that was required from grantees was a fairly standard research communications strategy in the grant application. PIPs emerged during the drafting of the Learning and Communications Strategy, with the emphasis on policy impact. The implications of this are discussed in section 6.1.
320. PACO comments on the PIPs cover the sections on context, risks, stakeholder engagement plan and learning indicators; and include suggestions for wider dissemination. Each set of comments takes approximately half a day of the Policy Advocacy Officer's time, which is a reasonable amount given the level of detail in the comments. Overall, the turnaround time for comments on the PIP appears to be relatively quick. 50% of PIs surveyed for this report indicated that responses were 'prompt' and 44% noted 'with minor delays'.
321. Progress on the PIPs is not a condition of tranche payment, though the team monitors grantees' reports for progress against the PIP, particularly focusing on pushing grantees to consider outcome rather than output indicators of policy impact. These might include asking grantees to report on actions arising from a workshop, not just the number of people attending it. Progress reports have only been received since August 2011 so the process is still new. Plans are in place to build a spreadsheet-based system for better oversight of progress against the PIPs. It is not unreasonable to ask grantees to report on actions arising, though care needs to be taken with the level of detail required. Given that implementing agencies often do not have the resources to track outcome-type indicators of impact, this may present significant challenges in future.
322. Further support is provided to grantees via Policy Influence Clinics, contracted out to ODI as part of the communications support project. Three clinics have been run to date (Cuernavaca, Rome and Colombo); a fourth is planned in Africa. These are two-day clinics for grantees and, as of 2012, their policy counterparts. Approximately 20 people have attended each clinic which involves a mix of lectures, networking and hands-on advice from communications and policy influence specialists. Up to five ODI consultants deliver each clinic, which means that the cost of each clinic is estimated to be in the region of £10,000. This seems expensive.
323. The evaluation team conducted a detailed review of the PIPs for the 21 sampled projects for this review. The reviews are in Annex 1. The quality of the PIPs has improved over time, both in terms of the line of argument about the relationship between the research project and policy, and the indicators of progress towards policy influence. However, PIPs are currently a half-way house between being a document intended to help researchers think about policy impact and something intended to help PACO monitor their progress. This raises particular challenges:
- The role of 'enlightenment' research does not seem to have been considered in the PIP process. This is public good research that percolates through the policy process over time, enlightening discussions and contributing to a body of evidence which will eventually influence policy (though how that happens cannot be predicted or subsequently attributed to a particular project). While policy relevant research will contribute to this

enlightenment function, not all research necessarily needs to have immediate policy influence. It is important to distinguish between policy influence and policy relevance.

Figure 17: Policy Advocacy, Communications and the grant process



- PIPs ask grantees to identify individuals at whom the research results will be targeted, as key policy influencers. But there may be legislative barriers preventing action, or other barriers to policy change which these individuals are unable to affect. Identifying individuals and then effectively holding them responsible for making change happen is a simplistic and risky strategy. This may be particularly true where 3ie funded projects are a part of a larger programme of work (such as a World Bank project). A detailed map of the policy process – done for each project – would help determine where the main points of leverage are for evidence in the policy process. However this is likely to be beyond the ability of most PIs and would incur extra costs. In addition, it raises the question of whether individual projects *should* have policy influence. This is addressed in detail in section 6.
- It is not clear who will monitor and report on policy impact after projects have ended. There is a tendency for researchers to make grand claims about longer-term impact which cannot be substantiated because nobody is monitoring it. Long-term assessments of impact would be easier if the individual projects summed to a coherent body of work, but

as is noted in the specialist reviews, this is generally not the case. In addition, PACO does not have a strategy or budget for long-term monitoring. This is an issue the current PIM project should address.

5.3.5 Summary

324. From the beginning, PACO has taken a strategic approach to its activities. The team has developed two strategy documents, and has then worked through them systematically. The small team has made effective use of outside help to develop its strategies and provide support to grantees. As the team has grown it has been able to take on many of the issues itself. There is a clear differentiation of responsibilities between the policy advocacy and communications aspects of PACO. Although detailed budgets are not currently available (and taking into account the comment in section 2.4.7 about the growth in the overall PACO budget and its relationship to 3ie's overall strategy), this division seems effective.
325. The website and e-newsletter are important platforms for external communications. They set up 3ie as a knowledge broker for IEs in general, rather than being simply a dissemination channel, as recommended in the KM strategy. Other external communications seem to be well received, with download statistics of the WP and SR series increasing in the past few months. The Journal of Development Effectiveness is rated highly by its publisher and has an exceptionally high ISI score for such a new publication. However, at an annual subscription rate of £78 per person or £232 per institution it could be questioned whether it is really contributing that much to building a culture of evidence within the 3ie target audience, especially the target audience in LMICs.
326. PACO has few responsibilities for internal communications, though communicating with members may become a higher priority if a new membership strategy is put in. If this happens, it may be more cost-effective to contract out production of the annual report to ensure that 3ie staff are able to spend more time liaising effectively with members.
327. The policy advocacy team puts a great deal of effort into liaising with grantees about improving their policy impact. Communications about the PIPs are prompt and generally thought by grantees to be helpful. Nevertheless, questions remain about the intense focus on policy impact, as outlined above. While both PACO strategy papers outlined techniques to improve impact and these have been followed consistently, the higher strategic question of whether an emphasis on policy impact is relevant to an organisation like 3ie was not discussed in either of them. This appears to have given rise to a degree of mission creep; with the emphasis having shifted from ensuring that projects practice good research communication and dissemination, to ensuring that this dissemination reaches policymakers, and to ensuring that policy is influenced.
328. There is no hard and fast line between good communications and policy influence, but the PIP process is currently neither one thing nor the other. PACO is responding to pressure to demonstrate that 3ie funded research has had an impact on pressing development questions. But PIPs as they stand do not do justice to the complex issue of policy influence. A thorough effort to influence policy would mean drawing up a detailed advocacy strategy using the full range of tools (stakeholder mapping, advocacy planning, and messaging and communication strategies). This would involve a great deal of work and is only appropriate where there is a large body of evidence from a variety of studies that have a clear message about an issue. As is argued in section 6, it is not appropriate for a study on a single issue to seek to influence policy directly, unless that study is part of an ongoing government policy or programme. The evaluation team's view is that while monitoring policy influence is necessary at the goal level (see section 6), the focus of effort should be just on good research communications.

5.4 Advancement and IE Services

329. 3ie's Advancement and Impact Evaluation Services unit is based in Washington. Advancement encompasses support to 3ie members and fundraising activities, especially in N. America. Until the recent appointment of Deputy Directors in Delhi and London, this work was mainly shared between the ED and the head of the Washington office, supported by the External Liaison Officer in Delhi.
330. There are five Impact Evaluation Services. Together they make up the major part of 3ie's strategy Component 4: Developing the capacity to promote, use and undertake Impact Evaluations.

5.4.1 Services for Evaluation Specialists

331. The first four of these services are resources for professional evaluators, researchers and IE specialists:
1. An Expert Roster of IE specialists: www.3ieimpact.org/en/evaluation/expert-roster
 2. An Impact Evaluation Database: www.3ieimpact.org/en/evidence/impact-evaluations
 3. A planned Registry of new Impact Evaluations.
 4. A Replication programme to encourage re-analysis and re-interpretation of IE datasets.

5.4.2 Services for Agencies and Governments

332. 3ie's fifth IE service is for implementing agencies and governments. Under the heading of Quality Assurance services, 3ie acts as a specialist adviser on impact evaluation. A current example is an assignment to act as external reviewers of Millennium Village evaluation for DFID.
333. 3ie staff also carry out a range of other activities, which overlap between Policy Advocacy and Quality Assurance: speaking at conferences, training courses, etc. Table 24 based on 3ie's Board report, summarises these activities for the six month period October 2011 to March 2012. The table covers a broad mix of activities. In the first quarter of 2012 these included: technical workshops and seminars with 15-30 participants (9); presentations and training for donor staff (7); participation in small donor discussion panels such as Bellagio (2); and, attendance at larger events, such as the African Evaluation Association conference in Accra. During this period there was also one Demand Generation Workshop, in the Philippines. By country, there were six events in the UK, five in the Netherlands, one in France and one in the USA. Among developing countries, there were four in the Philippines, three in India and one each in Ghana and China.

Table 24: 3ie QA, Training and Advocacy Activities

Final Table for the Board Report (October 2011 - March 2012)								
			Participants				Demand Generation Workshops	
3ie Staff Attending	All Events	3ie Events	Total	Policy makers/ Program managers	Developing Country	Institutions Represented	DGWs	Agencies Represented
ED	20	8	1040	761	755	382	9	34
Others	24	7	1532	542	370	489	1	50
Total	44	15	2572	1303	1125	871	10	84

334. The table underlines the considerable personal commitment the ED makes to this part of 3ie’s work: allowing for travel to the more distant events it might be as much as one month in six.
335. 3ie’s quality assurance practice has begun to extend into the coordination and management of donor-funded Impact Evaluation and Systematic Review programmes. Two important examples have been a joint AusAid-DFID-3ie call for Systematic Reviews, which was managed by the London Office, and a recent agreement to provide quality assurance and grant management support to a very large (\$20+ million) programme of impact evaluations of HIV Combination Prevention. This is being managed by the Washington office.

5.5 Planning, Budgeting, Monitoring and Reporting

336. 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. To illustrate this, Table 25 shows the Strategy objective statement for Component 1 and Table 26 shows the Detailed 2012 Workplan for Component 1. The last table in the set, Table 27, shows 3ie progress against the Workplan Indicators between 2009 and 2012. (This is also taken from the 2012 Workplan.)
337. When the 2011 Workplan was presented to the 4th Board meeting, there was extensive discussion and the Board made it clear it sought a stronger presentation: clear indicators, activities linked to strategic objectives, allocation of staff time, etc. 3ie’s Workplan 2012 was presented to the Board of Commissioners in October 2011, quite low on the agenda. The minutes report that the ED presented the Workplan and the Board adopted it, with *“some discussion regarding the quality and coverage of indicators, and their usefulness for monitoring 3ie performance. It was agreed that a review of monitoring indicators would be included in the process evaluation of 3ie.”* On this record it would seem that the Board were satisfied that 3ie had met the requirement they set out a year earlier.

Table 25: 3ie Specific Objectives and Targets for Component 1

Specific objective one						
To promote the generation of new evidence from impact evaluations that is relevant to policy and program design and implementation.						
What needs to be achieved (Key responsibility areas)	How it will be achieved (High level outputs)	Measurable Indicators at output level	Targets			
			2010	2011	2012	2013
Production of new evidence	Funding for primary studies	The number of completed and on-going studies	2/45	14/73	33/114	53/141
		The proportion of proposals funded which include developing country researchers and real country and/or NGO engagement	60%	65%	70%	75%
		Satisfactory rating in feedback from both successful and unsuccessful grant applicants	80%	85%	90%	90%
More and better production of evidence by other agencies	Provision of guidelines and facilitating guidance on specific studies	Satisfactory rating in feedback on our guidelines.	N.A.	80%	85%	90%
		Satisfactory rating in feedback on our quality assurance services.	N.A.	90%	95%	95%

Table 26: Detailed 2012 Workplan / Budget for 3ie Component 1

Strategy component	Key activities	Sub-activities	Budget (US\$ m)	Output indicators	Outcomes	Result		
Promoting the generation of new evidence from impact evaluations which is relevant to policy and program design and implementation.	Policy Windows	Management of on-going PW grants (PW1 and PW2)	3.123	PW study progress reports submitted	Research program oriented to policy priorities of funding agencies, and designed to needs of developing country policy makers	Basis laid for high-policy impact studies under PW supported studies		
		Launch and management of PW3	0.038	<i>No. of PW awards made</i>				
	Open Windows	Management and review of on-going OW grants (OW1 -3)	2.709	No. of completed and on-going studies (targets: 33/114)				
		Launch and management of OW4	2.149	Satisfactory rating in feedback from grant applicants (target: 90%)				
	Thematic window	Management of SPTW	1.304	<i>SPTW awards made</i>				
		Launch and management of TW2	0.012	<i>No. of TW2 awards made</i>				
		Launch and management of TW3	0.001	<i>No. of TW3 awards made</i>				
	Quality assurance services	Gates HIV/AIDS grant	0.064	Satisfactory rating in feedback on our quality assurance services (target: 90%)			Quality impact evaluations, with policy relevant designs, undertaken by other agencies	Development policy and programs informed by evidence
		Managing QA services	0.013					
		Responding to queries	0.020					
		Provision of QA services	0.041					
		Total Budget	9.474					

Note: Total Budget figure added by Evaluation Team

Table 27: Workplan Indicators, 2009-2012

Mission/component	Strategy indicators and targets	2009	2010	2011 provisional	2012 target
Mission: Increase the use of evidence	LMICs with on-going impact evaluations	Data to come from registry/ database			40
	Cases of demonstrable policy impact	Data from policy monitoring system			6
Component 1: new evidence from impact evaluations that is relevant to policy, program design, implementation	Primary studies – Completed	0	0	4	33
	On-going	23	53	70	114
	Proportion of funded studies with developing country researcher and/or implementing agency engagement				
	Grant applicant satisfaction (Applicant survey)	n.a.	87%	88%	90%
Positive feedback on 3ie guidelines (Applicant survey)		n.a.	95%	92%	85%
	Positive feedback on 3ie QA services	n.a.	n.a.		95%
	Component 2: existing evidence synthesized & disseminated in accessible, policy relevant form				
3ie Systematic Reviews – Completed		8	9	12	19
	On-going	2	4	7	8
Positive feedback on 3ie SR guidelines		n.a.	n.a.	n.a.	85
	Studies produced with 3ie management services	0	20	32	12
	Positive feedback on SR QA services	n.a.	n.a.	n.a.	85*
Component 3: Contributing to a culture of using evidence in development policy and implementation.	Primary studies published with clear, actionable policy recommendations	0	0	2	4
	Institutions taking actions to improve evaluation as a result of 3ie initiatives	n.a.	n.a.	n.a.	4
	Institutions & policy-makers influenced by 3ie-supported evidence	n.a.	n.a.	1	4
Component 4: Develop implementer & researcher capacity, especially in LMICs, to commission & undertake IEs, and to use the evidence.	Percentage of OW applications from LMICs (lead PI) (Targets: 60% and 30%)	n.a.	n.a.	n.a.	60% 30%
	Percentage of developing country governments which have received 3ie assistance reporting increased capacity to produce and use impact evaluations (Data to be collected as part of policy monitoring system)	n.a.	n.a.	n.a.	80%
Component 5: To build and operate 3ie capacity efficiently.	No. of staff in place (Target up to 32 by end 2012)	8	16	25	32
	Staff undertaking staff development activities	n.a.	100%	100%	100%
	No. of Board meetings	2	2	2	2
	Board approved workplan	0	0	1	1

338. It is a matter of judgement how much detail is required for effective planning and budgeting. In order to highlight the issues, the Evaluation Team have analysed the 2012 Workplan. From the minutes, it would appear that the Board did not review any of the financial and strategic decisions implicit in the workplan. And, on the material presented in the plan, it would have been difficult for them to do so. There are number of weaknesses, starting with the fact that no comparative figures are presented for budgeted and actual expenditure in 2011. The presentation is not helpful, combining a high-level, diagrammatic presentation of the allocation of resources across the five strategic components with a detailed budget allocation across 31 activities and 51 sub-activities, down to such immaterial items as \$1,000 on 'Guidelines'. The reader is left to his or her own devices to calculate links between the detail and the diagrammatic overview.²⁵
339. There are four 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. Secondly, 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. Thirdly, the only schedule of activities relates to the grant programmes. That apart, the workplan has no time dimension. Lastly, 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. If significant contingencies do arise, it would be better for management to make a formal application to the Board for a budget variation.
340. Specific new activities and budget shifts which the Board might have found it useful to consider include:
- A substantial increase over the 2011 allocation to supporting and expanding the membership base
 - The resource implications of taking on the \$20+ Gates HIV/AIDS grant
 - The reasons for carrying out some SRs in-house
 - \$407,000 allocated to Policy Monitoring activities
 - The fact that \$976,000 dollars for 'Institutional Arrangements: Related to HQ' included the large payment for the GDN hosting arrangement.
 - \$199,000 dollars for the IE Registry and the \$174,000 for Replication
341. The last item relates to two activities which have longer-term cost implications. As such, it would be better to present them to the Board for approval as separate projects, with a clear 'Business Case' justifying the commitment which is being undertaken.
342. 3ie has published two Annual Reports, for 2009 and 2010-11. These are produced to a high standard, with ample illustrations. Although they include details of the 3ie grant portfolio, membership and financial position, there is a strong advocacy/public relations slant. A number of case studies are presented seeking to emphasise the importance of rigorous impact evaluation. With most 3ie funded IEs still on-going, it is perhaps inevitable that some of the case studies are drawn from earlier work by other organisations: with a greater focus on the expected benefits of the interventions and the great need of the target populations than on the rigour of the theory of change and the strength of the evidence behind the interventions. These annual reports seem to be designed for a wider public audience; not as an integral part

²⁵ See note to Table 26 [IE where we had to add the total]

of a 3ie Planning/Budgeting /Reporting framework. Something rather different is required to provide clear management information and enable oversight by the Board and the membership.

343. During 2011, 3ie produced quarterly reports for Q1, Q3 and Q4.²⁶ These are short (5 – 7 page) statements of events during the quarter. The report for Q1 included some analysis of expenditures and grant disbursements and three attachments: a Quarterly Performance Monitoring Report using a more detailed version of the Workplan Indicators framework; a Quarterly Grant Monitoring Report; and, a Quarterly Financial Statement. The Q3 and Q4 reports are almost entirely narrative and seem to have had no attachments.
344. The decision to do away with the attachments is understandable. As already mentioned, the Financial Statements only present 3ie’s financial position on a cash basis. The Performance Monitoring Framework is not structured to provide useful management information on a quarterly cycle. It is based on the logframe model of indicators and targets, but it mixes practical indicators of 3ie output, grants completed/on-going etc., with rather subjective secondary indicators of stakeholder satisfaction with 3ie services and purely routine matters such as the occurrence of Board meetings. Few of these indicators are likely to change from one quarter to the next, and only the practical ones give a robust indication of 3ie’s progress. Proposals for a revised logframe, including a new set of indicators, are presented in section 6.2 dealing with the logframe. However, 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 must be a time-bound plan not just a list of activities and costs.
345. At the 7th meeting of the Board of Commissioners, 3ie presented a financial report for 2011. This was detailed and included analysis of variances against budget and commentary on how the variations had arisen. The most major of these were delays to the launch of IE windows. The report was thorough but it still lacked a clear statement of resource availability and analysis of how under- and over-spends related to delivery against time-bound milestones. Financial Projections 2012 – 2014 were also presented in some detail, with an explanations of how funds were allocated between grant and non-grant expenditure, and of non-core funding for ‘dedicated windows’. However, it was quite difficult to follow and gain a clear overview of the future funding position and of the critical balance between core funding and dedicated windows. In Section 2.8, the Evaluation Team have tried to show such an overview of the financial position and the availability of funds, using a cash flow analysis of future revenues against future commitments and separating funding available for 3ie’s core activities from those dedicated to specific programmes.

5.5.1 Assessment

346. 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. At the same time reporting against the agreed plan and budget should be the Board’s primary instrument of oversight and control.²⁷ 3ie does not yet have an integrated process of planning, budgeting and reporting which is fully effective for these purposes.

²⁶ The Evaluation Team has not had sight of a report for Q2.

²⁷ This is not stated in 3ie’s Accounting Policies and Procedures Manual, which emphasises the reporting requirements of individual 3ie donors. Only grants are expected to be managed ‘*to budget without material overrun*’.

347. 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, especially after the departure of the Deputy Executive Director in late 2011. With her replacement, in March, and the appointment of two more Deputy Directors, there should now be capacity to strengthen the process substantially.
348. The foundations for this are largely in place, in terms of financial control and time recording. The Evaluation Team has been informed that from 2013, departmental heads will prepare and be accountable for their unit budgets. This should allow 3ie's progress during 2013 to be clearly reported against Workplan and Budget, with variances between plan and budget reported at least six monthly, but preferably quarterly. (Given the nature of 3ie's work, monthly reporting would be unnecessary.)
349. Clear presentation is critical. As the Accounting Policies and Procedures Manual says, plans, budgets and reports must be presented in forms which management and the Board can use: *"tailored to the needs of the relevant non-finance manager, and easy to read and understand"*.
350. In this report, the Evaluation Team have endeavoured to offer some examples of this kind of presentation: the Cash flow Analysis shown in section 2.8; the discussion of costs in section 3.1, and the Table showing the Delivery Profile of the 3ie Impact Evaluation Portfolio 2.3.1.
351. This last example serves to make a 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 have heard of a certain disappointment among important stakeholders, even on the Board, that 3ie has yet to produce a large body of completed evaluations. As the table in section 2.3.1 shows, this disappointment is unrealistic. 2012 has always been the earliest date at which 3ie could be expected to deliver a substantial body of new evidence.

5.6 Financial Processes

352. 3ie's main financial transactions fall into five categories: grant expenditure; staff costs; professional services; travel and procurement. As processes, it is easiest to consider grant expenditures 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.

5.6.1 Grant Expenditures

353. As described above, grant payments are made in tranches against receipt of acceptable deliverables. With GDN support, 3ie check the grant applicant's budget. Where appropriate, greater detail is sought and budget amendments required 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.
354. 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.
355. A written Direct Costs Policy sets out what expenditures 3ie will and will not pay for:
1. Personnel Costs – Normally limited to the gross salary, excluding overheads/fringe payments. No payment is allowed to the agency implementing the intervention being evaluated.
 2. Travel – At economy or premium economy over 9 hours, with 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.
356. A separate Indirect Cost Guideline caps indirect costs at 10% of direct costs for government agencies, not-for-profits and universities, and 15% for NGOs and international organisations. This provision cascades to sub-grantees and sub-contractors, with the head grantee allowed 2% of the sub-grant direct cost as a Grants and Administration Charge.
357. The Evaluation Team have reviewed the records of nine OW grants to try to assess how these reasonable guidelines have been applied in practice. Although there may be gaps in the material seen, 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** – Co-funded by University of Chicago approximately 50:50, 3ie-GDN queried a survey cost of \$18 per 45 minute interview and 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 approximately 20%, in principle to fund a follow up survey against a completed baseline. However, the application form shows that over half of the 3ie contribution was used to cover unforeseen losses on the baseline resulting from the Asian flu outbreak. It is not clear why 3ie should have covered this risk. The total survey budget was over \$500,000: \$122 a head for a sample of 4,500. GDN-3ie asked for a more detailed breakdown. This showed a very large team, with 10 staff just for IT, but it was not challenged further. The grant application showed an overhead of 17.5%, which was reduced when requested, although the total sum remained the same. Although the lead PI was only charged at \$311 a day, four other academic staff were charged at \$1,000 a day. This was not questioned by GDN-3ie.
- OW 3.1229** – This application from the Mexican government National Institute for Public Health, sought \$130,000 for the re-analysis of a large existing survey, supported by a small qualitative survey. GDN-3ie asked for the 17.5% overhead rate to be reduced to 10% but accepted a reduction to 15%. The budget included three separate 5-day trips to conferences in the USA for three staff members. 3ie Evaluation Office queried this, but the question does not seem to have reached the applicant and the budget stood. A more significant question did not seem to have been raised: why a total of 478 staff days was needed for the re-analysis of existing data and a small, semi-structured survey of 129 interviews.
358. Some more general points emerged from the review. 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. It is understood that academics holding tenured positions at a US university should not charge for their time. On some grants, that is clearly the case. The PI's do not charge their time, or only the time spent in-country. On others grants, PI's are charged at near consultancy rates.
359. GDN-3ie checks on VfM seem to be focussed on the budget. Checks at the later stage seem less robust. Provided the EO and the EPA accept the deliverable, tranches are paid without any comparison of actual expenditure against budget. There are several cases where Tranche 2 and even Tranche 3 was been paid when the financial statement shows that Tranche 1 is almost entirely unspent. This leads to a bigger issue. On a number of grants, the Impact Evaluation as implemented is significantly different from the original proposal. Although 3ie was only a 20% co-funder, OW 1.43 is an example. The proposal was for a survey of 4,500

households. The final reported number of completed surveys was 2,843. GDN aims to balance any difference between actual expenditure and budget in the final tranche, and one grantee has returned unspent funds. The Evaluation Team feels this final-balancing approach may leave it too late to ensure a clear settlement, in particular on co-financed studies, where savings from underspends must be allocated between the co-funders.

360. It emphasised again that only nine grants have been reviewed and 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. Closer attention to Value for Money, as delivered, in addition to as proposed, may be particularly important; with the option of early termination if the possibility of an outcome commensurate with the cost becomes too small. Shifting the balance of the tranche payments, and requiring grantees to tie them 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 should be closely involved in the assessment of the VfM of grant proposals and grant execution. Their judgement on whether the effort is justified by the scope of work is fundamental.

5.6.2 Non-grant Expenditures

361. 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%).
362. For staff costs, 3ie's processes are largely determined by those of the two hosting organisations: GDN and LSHTM. As a small newly established organisation, salary structures have perhaps developed ad-hoc, especially where staff have started work on contract before transferring to a staff post with GDN or the London School. 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.
363. The 3ie consultancy budget is split between support to the grant management process and more ad hoc requirements. For the first, 3ie uses a large panel of specialists as External Reviewers/External Peer Advisers paid a standard fee of \$200 per review. Up to OW3, 3ie had used 82 different specialists, who received an average payment of \$1,800. The reviewer with the most work received just under \$10,000. For a simple review of high quality applications, a reviewer might complete three, at most four reviews a day. However, it is the Evaluation Team's experience that it is difficult to keep this rate up, and the record of the reviewer's comments show that they have put a substantial effort into many of the reviews. As will be discussed later, the role of the EPA is particularly important. Overall, this service has given 3ie good value.
364. 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.
365. 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.

5.6.3 Hosting Agreements

366. At \$868,000 in 2011 The GDN hosting fee was 3ie’s second largest non-grant cost, after salaries and benefits. The fee includes a base element of \$600,000 and a percentage management charge on against the volume of 3ie grant disbursements above \$4 million.²⁸
367. The base fee 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. The Evaluation Team have not seen a breakdown of GDN’s costs, but the MoU states that the rent element is INR 9,216,00 (\$200,000 at the time the MoU was signed, \$165,000 now). Using some rule of thumb figures a reasonable estimate of how much it costs GDN to host 3ie might be:

	\$
Rent	200,000
4 staff @ \$15,000	60,000
Fixtures, fittings depreciation	20,000
Travel office @ 10% of spend	<u>43,000</u>
Total	323,000
Management fee on grants < \$4million	<u>277,000</u>

368. At grant disbursement of \$4 million the fee element calculated here is equivalent to 7%, the same as the rate set in the MoU for disbursements between \$4 million and \$6 million.
369. In value for money terms, two elements of the GDN agreement merit attention: the rent, and the percentage management fee. The rent is for a large, well-appointed modern office in a mid to high level area of Delhi. The Evaluation Team are not qualified to judge on how the rent compares with market rates or how important it is that 3ie has a prestige address.
370. In Q1 2012, 3ie made 22 disbursements on OW grants, two on PW grants and seven on SR grants: 31 in total. 51 small payments were also made to reviewers. 81 transactions in a quarter equates to about six per week. Given that 3ie is paying a full cost for office rent, etc., that the majority of the work is handled by the 3ie PO, and that GDN has one of the four 3ie-dedicated staff working in its finance office, it is difficult to see how this generates a burden on the GDN overhead equivalent to 7% of the grants disbursed.
371. 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.²⁹ In other respects the arrangement seems quite cost-effective, especially if LSHTM is carrying the cost of pension and insurance cover for the staff.

5.6.4 Procurement

372. 3ie’s Accounting Manual sets out three levels of purchase: up to \$500; \$501 to \$50,000; and over \$50,000. For the second level, a relatively simple three-quotes process is laid down. For purchases over \$50,000, the only extra requirement is second approval from the ED. No additional procedure is laid down for purchases over \$200,000, for which the Board Chairman’s approval is also required according to the Authorisation Matrix.

²⁸ The base element has been increased to \$650,000 for FY 2011/12.

²⁹ This is assuming the stated salaries in the MoU do not include a margin to LSHTM

373. Grant management aside, procurement is not a major function for 3ie. The simple procedures laid down are adequate for most requirements. However, 3ie has awarded three substantial contracts, well over \$200,000: the ODI contract for Policy Influence Monitoring(PIM) (\$400,000), the Praxis contract for the GMS (\$400,000), and, recently, a contract award to the Rand Corporation to develop an IE Registry (\$790,000).
374. The PIM contract was awarded through a competitive tendering process against a detailed RfP. This resulted in six bids being received, all of which were reviewed by a four-person expert panel with a wide range of expertise (policy influence and evaluation) which included two external advisers. The expert panel scored each proposal on methodology, organisational capacity, regional capacity and professional experience and produced a clear winner. The evaluation team has reviewed the RfP, the detailed evaluation sheet produced by the expert panel and the final proposal. This was a competitive and well-run tendering process.
375. The Praxis and RAND contracts were awarded through a competitive process against a written RfP. In both cases, three proposals were reviewed by small panel of 3ie staff with one outside adviser and formal selection memo submitted to the ED and the Board for approval. The difficulties with the Praxis contract are set out in Section 5.2.3 above, along with the Evaluation Team’s assessment. The Team has also briefly reviewed the RfP and Statement of Work for the RAND contract. The specification laid out in the RfP is less detailed than the one given to the GMS bidders, and the Statement of Work submitted by RAND reads as an academic discussion of the principles of a registry. There is no clear statement of exactly what will be delivered. Instead it makes open ended statements of possibility: e.g. *“One standard that needs to be explored is the Data Documentation Initiative (DDI). Using standards enables the platform to interact more easily with other public registries (for example using XML Document Type Definition (DTD).”* In IT contracts, open-ended statements like this could rapidly turn into open-ended financial commitments.
376. 3ie have not felt able to share the RAND contract with the Evaluation Team. It is possible that includes something much more specific. Absent that, the Team feels there is a risk of similar problems to those with the GMS. It would also note that the external adviser on 3ie’s review panel for the Registry contract was an evaluation specialist. As already emphasised with reference Praxis, it is essential to have capable IT advice when commissioning IT projects.

5.7 Management Processes and Style

377. 3ie has been a very reflective organisation with a strong culture of continuous learning. The steady development of the Grant Management process is just one example of this. Apart from regular surveys of grant applicants, those attending its conferences and other stakeholders, 3ie has commissioned three reviews of its internal operations:
- May 2011 – A Rapid Institutional Review
 - July 2011 – London Office Institutional Review
 - August 2011 - A Culture Climate and Awareness Survey (CCAS)
378. As a small organisation, the Executive Director position has been fundamental. To win credibility for 3ie as a centre of excellence for Impact Evaluation, the ED has had to demonstrate the highest possible level of technical skill and authority, and carry out a heavy programme of advocacy and public presentations, in a wide range of fora from high-level international conferences to demand generation workshops in developing countries. At the same time he and his deputy have put in place the systems to manage a substantial

programme of grants and supporting activities. The ED's performance has been a focus of attention, culminating in a 360° performance review commissioned by the Board in July 2011. The CCAS formed part of this review.

379. The concentration of reviews in mid-2011 suggests a degree of concern about 3ie's organisation and management. Since that time, partly in response to these reports, a number of organisational changes have been made, most importantly the appointment of four Deputy Directors. The Process Evaluation is not designed to repeat the institutional reviews. Instead, this section presents a brief overview of the conclusions of the earlier studies, with limited comments on how changes in the last year have addressed any concerns identified. Where the reviews make points which are covered elsewhere in the Process Evaluation, they are not repeated here.

5.7.1 The Institutional Reviews

380. The Rapid Institutional Review (RIR) and the London Office Review were carried out by the same consultant, so they can be considered as one exercise. The two studies were linked to 3ie's 2011 Annual Staff Retreat, at which the consultant acted as facilitator. The opening qualification to the RIR should be noted: that it is "intentionally short and schematic, given the insubstantial and impressionistic nature of the process followed to compile it."

381. The RIR 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"

382. 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.

383. The report on the London office describes a particularly severe case of 3ie's issues with over-commitment: *"workload is excessive and needs to be better organised with a clearer eye on the overall intended results [work] would benefit tremendously from being better framed and planned. Of particular concern is the tendency of the office, consistent with 3ie practice, to accept additional responsibilities and assignments, without ensuring that already over-stretched capacity is supplemented. As a result, performance on standing and important objectives is compromised in favour of ad hoc, less important matters."* This longer quote is included by way of highlighting the point about the need for annual workplans which set out clear and time bound operational indicators for each activity. That is what ensures that staff are clear on which objectives are 'standing and important', i.e. those objectives which have been approved by the Board of Commissioners.

384. The report goes on to note a lack of differentiation in the responsibilities of London staff and specialisation in their activities. Presumably in response to these criticisms, the London office now has an annual workplan, running from July 2011 to June 2012. However, it remains quite unspecific as to targets and allocation of staff responsibilities. A typical item is shown below. It lacks any indicators which can be described as SMART and the staff time allocation is for the whole SR Team. Who will actually do the work is undecided.

Table 28: Sample of SR Team Workplan

Activity	Quarter 1 July-Sept		Quarter 2 Oct-Dec	
	Target measures	Time allocation (%)	Target measures	Time allocation (%)
4. Campbell IDCG management	IDCG: WASH, FFS, CCTs and governance titles published on Campbell IDCG webpage; protocols submitted to Campbell for peer review; external author teams receive support on titles, protocols, reviews; Col policy produced; co-registration documents produced/agreed; preparation for annual Campbell symposium; preparation for MAER-net colloquium.	25%	IDCG: author teams receive support on titles, protocols, reviews; preparation for Cochrane colloquium.	15%

385. Although only five international evaluation specialists contributed directly, the RIR’s note on Views of 3ie in the Evaluation Community is a clear summary of three quite divergent positions:

- *“3ie is seen as an authoritative voice on impact evaluation, ... having demonstrated a commitment to building North-South research partnerships and capacity. ...”*
- *“3ie is something of a disappointment and a wasted opportunity to build capacity and change the evaluation agenda. ... despite its inputs on mixed methods, the Initiative is still largely focussed on quantitative methods under a theory of change that one big study can come up with THE answer”* Summed up in the report as an impression *“that 3ie does not walk its own talk.”*
- *“3ie is a self-serving club for funding a small number of Northern researchers’ particular research agendas which are not connected to policy needs at all.”*

5.7.2 The Leadership and Organizational Performance Review

386. This review had three parts: a 360° review of the ED, the Culture and Climate Survey of 3ie staff, and a survey of 3ie Members.

387. The 360° review compared the ED’s self-rating on a range of Functions/Leadership practices with the rating of the Executive Committee of the 3e Board (B1), the rating of other Board members (B2), a group of peers in the IE community (P) and the ED’s Direct Reports (DR). The Evaluation Team do not propose to review the results in any detail. In general they show 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 demonstrate a 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.

388. The CCAS is similarly supportive. In most respects, 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. It was, however, disappointing that the lowest score of all was in response to the statement 'I see myself working in this organization 3 years from now': just 46% in agreement.
389. 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.
390. 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, member representatives felt there was not enough information to gauge the *quality* of proposals, a top priority during 3ie's start-up. They were also uncertain that the diversity of 3ie's portfolio was the most effective way to build an aggregate base of evidence in the strategic areas.

5.7.3 Assessment

391. The set of reviews carried out in mid-2011 show a consistent picture, one that is well summed in the words of the members interviewed for the Leadership and Organisational Performance Review. They set out a clear view. It starts from recognition of 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 more 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.
392. 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.
393. 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 ensure that such opportunities are not undertaken unless adequate spare capacity is available.

6 Relevance and Effectiveness

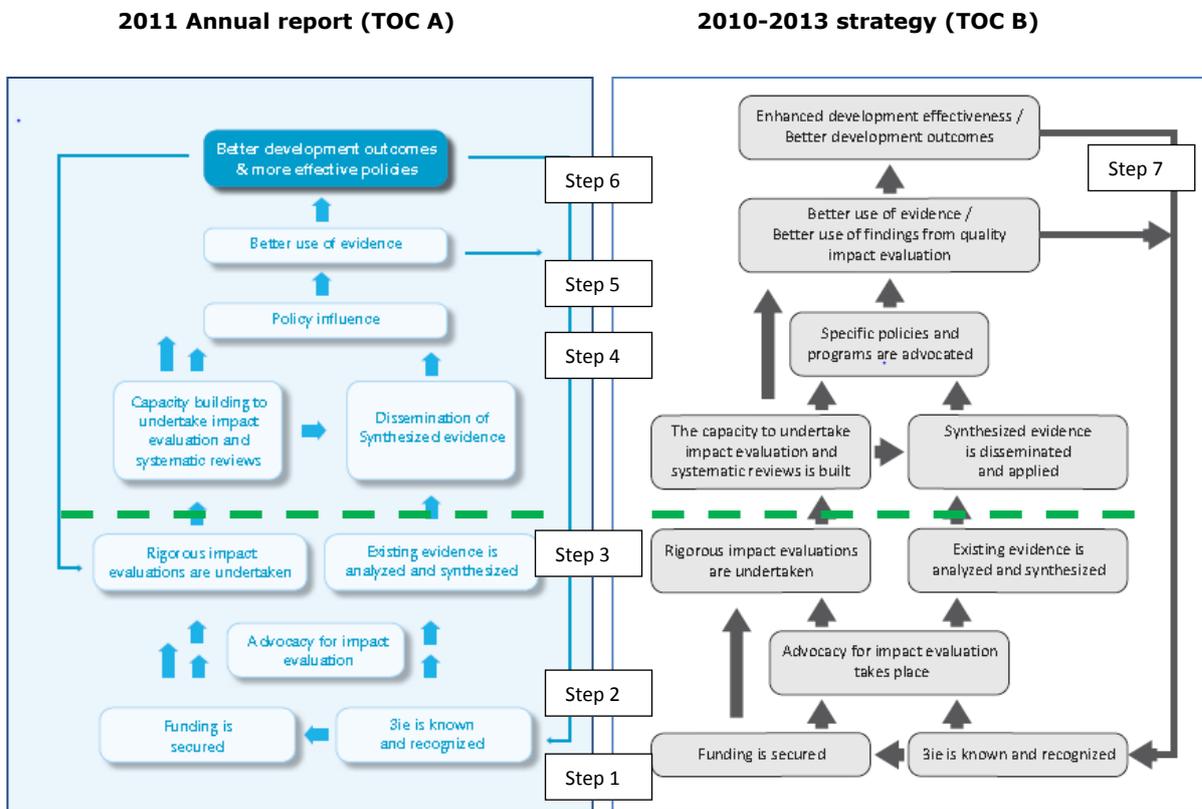
394. In this Chapter, the Evaluation considers two different aspects of 3ie’s Relevance and Effectiveness: the Theory of Change, with its the underlying logical framework, and the Potential Impact of 3ie grants, as the best immediate measure of the relevance of 3ie’s efforts to generate rigorous, policy relevant evidence.

6.1 The Theory of Change

6.1.1 3ie’s Current Theory of Change

395. 3ie documents present two slightly different versions of the Theory of Change (ToC). These are set out in Figure 18.

Figure 18: The 3ie Theory of Change



396. The ToC is simple and more or less linear. Below the dashed lines the logic is the same for both versions: that advocating for impact evaluations leads to rigorous IEs and synthetic reviews being conducted, thence to their dissemination and to increased capacity to undertake further research. Above the dashed lines, the two theories divide. In ToC A, dissemination and capacity building leads to improved ‘policy influence’, which is not defined but assumed to encourage the better use of evidence, leading to better development outcomes and more effective policies. Feedback loops improve 3ie’s recognition and the undertaking of more rigorous IEs. While there is no exploration of the assumptions which underpin it, the text of the annual report gives a set of impact measures for 3ie’s advocacy and capacity building work, dissemination and publication, improvements in the culture of evaluation, and cases in which 3ie has had a

demonstrable contribution to project, programme or policy design. A final indicator, at the development outcomes level, is the number of LICs and MICs with on-going IEs.

397. In ToC B, there is a greater emphasis on advocacy. The key difference is that specific policies and programmes will be advocated and that this (combined with capacity building to conduct IEs) will lead to the better use of evidence in general and IE evidence in particular. The final outcome mentions development effectiveness and development outcomes, but does not specifically link this to more effective (generic) policies. There are no accompanying measures of impact which would help anchor understanding of what will be achieved.
398. This lack of clarity in the top end of the ToC goes back to early 3ie documents, which do not specify how influence on policies and programmes is expected to come about. It is possible to deduce elements of this part of the theory of change from different papers, although they still do not add up to a clear logic for this part of the ToC:

Document	Implicit aspects of a theory of change
3WEL (Table 2, p 36)	Impact on policy is inferred from a combination of issues: internally-commissioned studies gaining external legitimacy through the work of ‘the council’ ³⁰ ; improving incentives for staff to conduct IEs, focusing the IEs of other organisations on their own needs, swapping experiences with other countries and being able to access high quality research/researchers more easily. It is implicit that greater involvement of social ministries in Impact Evaluation will lead to a broad improvement in the use of rigorous IEs. The paper calls for advocating for legislation and policies, but it is not clear whether this means for specific social programmes or for a more general advocacy strategy to support the production and application of knowledge.
3ie founding document	Does not contain any analysis that could be described as a theory of change, though section 10 notes that studies likely to provide evidence useful to policy decisions in LICs and MICs are those which are likely to be replicable and/or scaled up, which guide efforts to improve management or redirect the intervention towards more successful approaches, or which help judge the generalisability of IE findings.
3ie Learning strategy for policy impact	Sets out a series of objectives with accompanying activities that will be undertaken to ensure that 3ie-funded grants influence policy, without defining policy impact or how it will be measured. Unlike 3WEL, the implicit definition of impact does not reflect demand from member organisations. Instead it emphasises evidence-based campaigns, and positioning 3ie as a trusted broker of evidence from IEs and SRs.
3ie/ODI Knowledge management strategy	Identifies a narrow 3ie definition of policy and policy influencing. Suggests a focus on policy influencers in donor agencies or international thematic bodies will be more effective. No definition of what impact on these stakeholders might look like, or of a broader theory of change.

³⁰ 3WEL refers to ‘the council’ as the entity which would subsequently evolve into 3ie.

6.1.2 Analysis

399. The theory of change can be assessed from the bottom up in seven steps, as shown in the figure above.

Step 1. 3ie is known and recognised ↔ funding is secured

400. As discussed in Chapter 6, 3ie is well-known in the wider evaluation community. It is less clear what it is known for. This is a first weakness in the ToC. It is perhaps implicit that it be known as an advocate for IE. From the original evaluation gap discussion, and the ToC figure, it seems that means an advocate for greater funding for IEs. However, 3ie is in fact now known mostly as a funder of IEs and SRs, and as an advocate for experimental/quasi experimental evaluation techniques.

401. As it stands, the link between recognition and securing funding is weak. Funding for IEs is not dependent on the number of people or organisations familiar with 3ie. However, it seems clear that this part of the ToC can be regarded as completed. Whether thanks to 3ie or not, ample funding is now available for IEs.

402. For reasons which will be discussed in much more detail in the next section, the Evaluation Team believe this part of the ToC should be revised to read:

“3ie is recognised as centre of excellence for Impact Evaluation”

With the link to funding removed.

Step 2. Recognition plus funding facilitates Advocacy for IE ↔ IEs being undertaken and evidence being synthesised and analysed

403. 3ie has been a strong advocate for impact evaluation; through its website, publications, newsletters, conferences and seminar series. It has drawn in funding not just for its core operations but also for dedicated thematic windows, and systematic review programmes. It has provided training on the principles of IE to donor staff and others. However, it is arguable that this may have had little impact on the volume of IEs and SRs undertaken. Donors are committed to the generation of evidence. Without 3ie, they would have found other channels.

404. There is one exception to this: 3ie’s effort, through its Policy Windows, to stimulate IEs which respond directly to LMIC policymaker demand. This is something which is not matched by other programmes. However, as it stands this important activity does not really fit into the ToC. As the upward flow underlines, this is essentially a supply-side theory. The least awkward way to make the demand side effort explicit, might be to state that advocacy for IE creates demand from country governments.

Step 3. Rigorous IEs plus systematic reviews ↔ Capacity [LMIC?] to undertake IEs/SRs AND ↔ Dissemination of Evidence.

405. This Step presents two different channels through which the TOC expects impact evaluation to have influence on policy. The first is implicit. It appears to be expected that the involvement of LMIC researchers in IE work, and the development of their capacity will give the IEs they work on greater policy influence. This is logical enough and 3ie has made it a clear part of its policy that LMIC researchers be involved in implementing IEs to the greatest possible extent, and that they be given opportunities for capacity building. As discussed below, this has been largely limited to on-the-job training, but is not necessarily the worse for that. However, there is no evidence of this leading to policy influence. Absent a strategy to give country researchers the lead in the PIP process, and probably some budget to go with it, this may be a relatively weak link.

406. As presented in the ToC, dissemination is the only other channel to policy influence. In practice this conceals 3ie's very considerable efforts to promote the policy influence of the evidence it generates. Those are discussed as the next step. 3ie has put a great deal of effort into dissemination of synthesised evidence via its various publications and the website: the vast majority of survey respondents found both to be useful. However, it is well established that dissemination is only a first step towards real influence and policy change.

Step 4 *Capacity AND Dissemination ↔ Policy influence / specific policies and programmes are advocated*

407. The ToC weakens significantly at this stage, even when 3ie's policy influence efforts are taken as being implicit in the theory. There are several areas of weakness:

408. First, 3ie funds two very different types of project, evaluations of on-going government programmes and proof-of-concept trials. Each will be treated differently in policymaking processes. On-going government programmes will have organisations or groups with identified responsibilities for implementation. If the IE comes up with evidence that change is necessary, it will be relatively clear whose responsibility it is to make that change. The same is not true for proof of concept studies, where issues such as scalability, distributional aspects, responsibility for implementation and cost have yet to be addressed or budgets allocated. The evidence from these types of project will feed into various working groups, technical committees and other planning processes where the project's evidence will only be a part of a wider package of evidence and analysis which – depending on other government priorities – may or may not see the light of day. Policy is not made on the basis of a single proof-of-concept study.

409. Regardless of the type of study, 3ie requires researchers to work to a standard format Policy Influence Plan (except for PW grantees). These are 3ie's main vehicle for achieving policy impact. They ask grantees to consider the wider policy context and the risks inherent in that context, and to make a stakeholder and policymaker analysis. This is used to identify individuals who have the ability to influence policy, to develop a stakeholder engagement plan, and draw up a set of indicators for monitoring the impact on policy.

410. Identifying individuals who can influence policy is a central part of this process. There are two problems with this. First, it may encourage grantees to identify individuals who they believe are more likely to take their results on board, leading to a selection bias which discourages rigorous analysis of the study's policy relevance. Second, it presumes that these individuals welcome contact with the project and are willing to provide feedback on the project when tranche payments are due. For high-level policymakers this is a substantial ask that can be counterproductive to good relationships with project staff.

411. 3ie's stated intention is to be an independent advocate for robust evidence. However in essence, 3ie's interpretation of this section of the ToC presumes that IE researchers should have the remit to use their evidence to advocate for policy change. There is a danger that this approach will cause mission creep, as shown in the diagram overleaf.

Figure 19: Mission creep in advocating for robust evidence in policy processes



412. PIs who responded to the survey had a mixed reaction to the issue of policy advocacy. 22 out of 36 were comfortable taking on advocating for policy change based on their results. 14 were not – the majority of these would have preferred working with a policy-facing organisation to get their message across.

413. It is too early to judge whether 3ie’s work will enhance the policy influence of IE evidence. It is perhaps not very different from what other programmes encourage. In the end, the policy impact of IE evidence will depend more on the quality of that evidence, and the extent to which it meets current policy needs, than it does on a structured approach to dissemination and influence. In other words, quality and policy relevance are the real guarantee of policy influence.

Step 5 Policy influence ↔ Better use of evidence

414. This step seems redundant. If anything better use of evidence should be considered as an integral part, or even precede, policy influence. Better use of evidence can only come from some form of capacity building or advocacy for policy formation, which seems outwith 3ie’s mandate.

Steps 6 & 7 Better use of evidence ↔ More effective policies or Better development outcomes

415. Following on from the preceding, a more logical wording for this is “More and better evidence leads to more effective policies” The bottom half of the Theory of Change is clearly about stimulating the production of more evidence of a higher quality. To interject use of evidence into the process only serves to undermine the clarity of what should be a simple theory: that development outcomes can only improve if there is more, and more accurate evidence available.

6.1.3 Summary

416. 3ie’s Theory of Change is logical and straightforward but imprecise in areas. The following assumptions seem implicit, either in the theory itself, or in 3ie’s interpretation of it:

- 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 given: that the research community, of its own accord, will generate enough policy relevant proposals that it will only be necessary to select which are the most policy relevant,

- Somewhat illogically, given the previous proposition, that policy influence requires a major effort which must be built into the research from day one.

417. The Evaluation Team believe there are problems with each of these assumptions. For the first, 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 sets out reason to suggest that the quality dimension cannot be taken as given. There is an urgent need for IEs to be more rigorous. Researchers are facing significant difficulties translating experimental/quasi experimental techniques to the development context. Policy relevance is also quite variable. In a number of sectors 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*.
418. A consultation on the October 2009 version of the 3ie strategy document³¹ showed that 3ie staff, counterparts, Board members and donors were all pushing strongly for policy impact. Board members were particularly keen to see policy impact, though there were concerns about how to measure it and whether the strategy had a sufficiently strong Theory of Change. Subsequent discussion at that Board meeting³² agreed that while the vision statement would be aspirational (not something measurable and attributable to 3ie), the mission statement should be pitched at the level of influencing policy. These sentiments have been repeated at subsequent Board meetings, as has the idea that 3ie can advocate for robust evidence but not for a particular policy position. As shown above, it can be difficult for individual projects to draw this line.
419. 3ie has sought ODI advice on this upper level of the Theory of Change. However, neither party seems to have taken a step back to question the validity of the last assumption: that a targeted policy influence effort should be a requirement of every grant. It is surprising that the initial ODI consultancy did not explicitly explore it. Similarly the current PIM project is to discuss the individual theories of change with each project it monitors, based on the useful framework developed by the 3ie Working Paper by Weyrauch & Diaz Langou (WP12). However it does not appear to be stepping back to consider the wider issue of whether 3ie should be focusing on policy influence as a matter of course for all projects.
420. In the same way that it is possible to slip sideways from advocating for robust evidence to advocating for a particular policy position, the emphasis on policy influence could be the result of a similar sideways slip from ‘good research communication’ via ‘good research communication to policymakers’ to ‘influence policymakers with the results of research’, compounded by donor pressure to demonstrate the impact of aid spending on sustainable policy change. Again, this introduces a potential mission creep:

Figure 20: Mission creep between good research communications and policy influence.

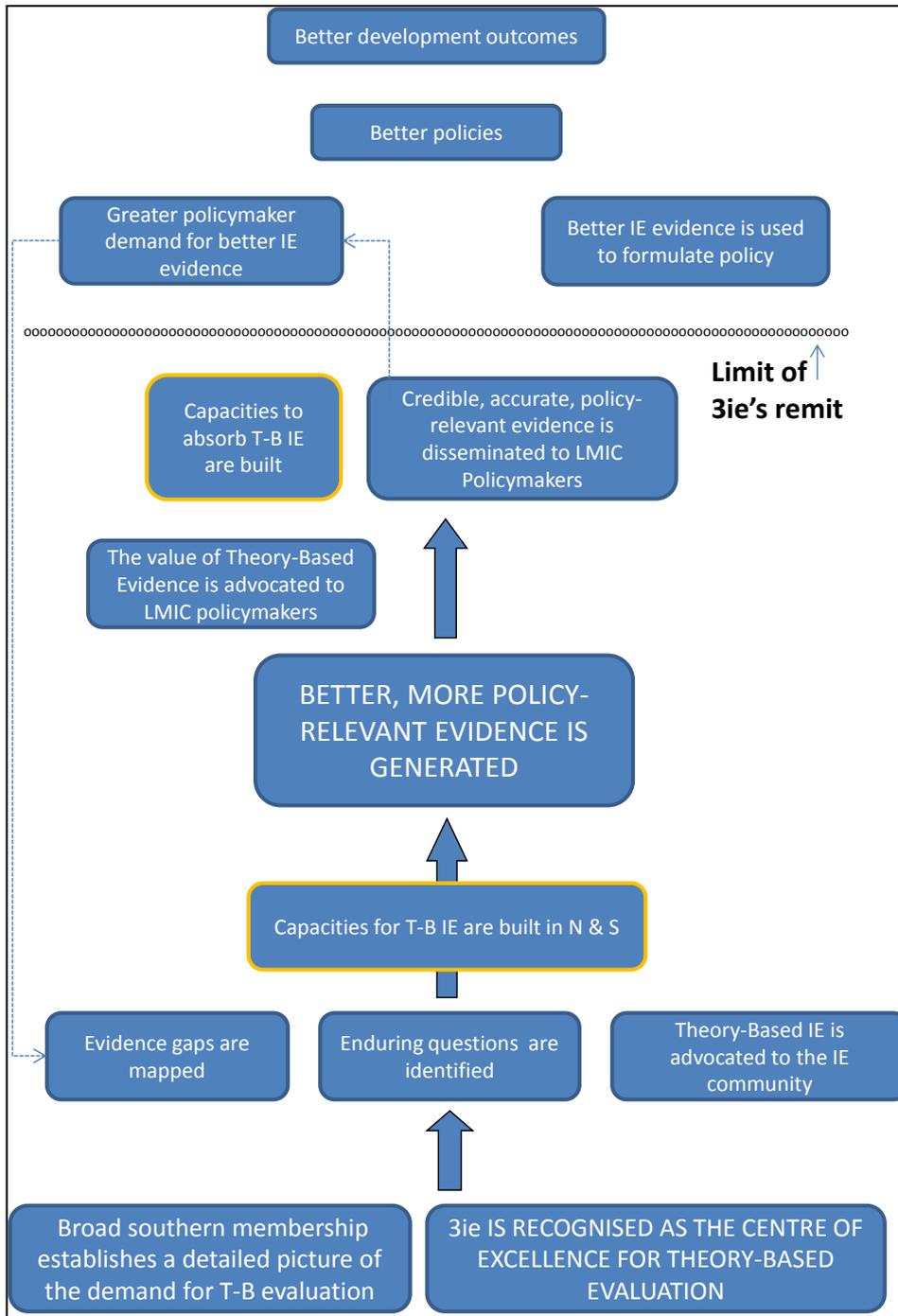


³¹ D Fraser presentation to the Jan 2010 Board meeting

³² See minutes of the Jan 2010 Board meeting, Annex B

To conclude this review, the Evaluation Team present an alternative Theory of Change in Figure 21.

Figure 21: Suggested Theory of Change



Note: boxes ringed in orange are not entirely 3ie’s responsibility

421. This revised theory of change starts from the proposition that there is an urgent need for greater rigour in IEs carried out in developing countries, and that rigour does not come solely from the application of statistical technique. As a technical matter that can be taken as a given. Instead, it is 3ie’s Theory-Based approach which is the true gold standard for impact evaluation in developing countries.

422. While 3ie is not wholly responsible for ensuring that there is sufficient capacity building in both north and south to carry out theory-based impact evaluations, it has a clear role to play in working with a broad southern membership to identify their common or enduring questions and map where the evidence gaps lie. In also advocating theory-based IE to the IE community, 3ie thus builds the potential to generate better (more rigorous, more theory-based) and more policy relevant (responding to its membership) evidence which it disseminates to LMIC policymakers. And its engagement with members will, as a matter of course, advocate theory-based IE to policymakers. The combination of these two activities will lead to more policymakers demanding, and using, this high-quality evidence. The greater the demand from policymakers, the more 3ie will be able to map the evidence gaps and feed this information through into its grant commissioning processes.
423. It is important to be clear about the limit of 3ie's remit. While it is important to monitor above the line for policy influence, the evaluation team have cautioned elsewhere about the potential for over-emphasising policy impact of 3ie grants. This revised theory of change avoids the potential for mission creep, showing that it is possible to retain the potential for policy impact via a clear focus on the production of rigorous, theory-based impact evaluations. This is fleshed out in the next section.

6.2 The Logical Framework

424. Two logical frameworks have been used during 3ie's lifetime: one in DFID's final project memorandum and the other in the proposal to the BMGF (neither document is dated). The BMGF logical framework was used in the recent review. The Evaluation Team understand it was felt that while the objectives were still relevant, the output indicators were weak. The discussion in the previous section would suggest a radical revision. However, it is not appropriate for a Process Evaluation to undertake that. Any major re-drafting of the logframe should be led by the 3ie team and involve all major stakeholders. This section seeks merely to suggest ways the current logframe might be strengthened and provide a clearer framework for monitoring 3ie's progress against robust indicators.

6.2.1 Programme Logic

425. Table 29 presents a draft for a logical framework revised on that basis, taking the BMGF version as the starting point. Items shown in italics are suggested additions. Those in square brackets have been moved up the logframe. A strikethrough indicates the item does not fit the in logic. For clarity, the five objectives have been placed on separate boxes. The proposed logic changes are as follows:
1. A minor but significant change to the Goal statement does follow on from the discussion in the preceding section. The Evaluation Team believe the simple reversal of the phrase 'better use of evidence' to 'the use of better evidence' substantially strengthens the logic. 3ie's main purpose is to support the production of more, better evidence; not to ensure that evidence is used in a better way.
 2. The one major revision has been to introduce a Purpose statement. This helps to clarify where 3ie's responsibility stops. It is standard practice that the Purpose is the highest level on which a programme can have a direct impact. This change indicates that it is 3ie's Purpose to deliver IE evidence which influences policy and implementation. Whether or not that influence leads to better development outcomes at the Goal level is a fundamental assumption of the Theory of Change. But it is not something that can be directly attributed to 3ie.
 3. In the current logframe indicators of policy influence are placed at the Objective (or Output level). These have been moved up to the new Purpose level, where they fit the logic better:

‘programme designs changed by IE evidence’, ‘institutions influenced by IE evidence’, etc. Two critical assumptions have also been moved up: about the political limits on reform, and the possibility that IEs will generate useful evidence.

4. One new Purpose indicator is proposed, to bring the concept of Enduring Development Questions into the logframe. It is suggested that answers to these questions will have the greatest likelihood of influencing policy, especially among the donor community.
5. One new Purpose assumption is proposed, although it is just a re-wording of the existing assumption about the political limits to reform.
6. 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 assumption. In other words, if it fails, then 3ie should never have been started:

KILLER ASSUMPTIONS – ‘Impact studies have the capacity to identify welfare improving policy changes.’ If this is not true, then the whole concept of the Evaluation Gap is false.

- ‘3ie will attract sufficient funding’. Without funding, 3ie would never have started. This is also an internal assumption, in that it is up to 3ie to persuade funders that it is worth supporting.

INTERNAL ASSUMPTIONS – ‘Studies supported by 3ie will appeal to policymakers’. This duplicates the Purpose assumption about the possibility of influencing policy. However, it is also something that 3ie should be able to influence directly: through the relevance of the studies, and the quality of its communication.

- ‘Funding available for capacity building’. This is a management decision for the 3ie Board, not an assumption.

6.2.2 Indicators

426. As discussed in Section 5.5, it is recommended that 3ie introduce a lower level of operational indicators, tied to the planning and budgeting cycle, and that those are used to monitor 3ie’s progress, quarterly or half-yearly. It is a question of judgement how much of this operational detail should be brought up into the logical framework. In the draft proposed here the aim has been to provide a more concrete set of indicators, with less dependence on secondary, subjective measures of 3ie’s performance such as stakeholder satisfaction, but without excessive operational detail.
427. **Objectives 1 and 2** – The number and quality of completed IEs and SRs must be the fundamental indicator for this objective. The proposed revisions add further quality dimensions to that. Although the management of grant windows is an operational matter, it would be useful to include summary indicators of that progress in that respect at this level.
428. **Objective 3** – Summary operational indicators have been added: on events, publications, etc. A somewhat open indicator on ‘taking actions to improve the culture of evaluation’ has been revised to focus on specific actions taken.
429. **Objective 4** – Indicators on 3ie’s Policy Window and LMIC government initiatives have been made more specific. The link to Objective 1 through the involvement of LMIC researchers has been simplified. An indicator on PPGs is left in place, but bracketed to mark the fact that this initiative has now stopped.
430. **Objective 5** – This has been strengthened in various ways, starting from the addition of a two key indicators of institutional strength: the level of core funding available to support 3ie’s work, and the Membership. Two weak existing indicators – Board meetings and workplans approved – are hardened as: ‘No of Board actions completed/uncompleted’, and ‘Workplan objectives delivered.’

Table 29: Draft for a 3ie Logical Framework

Note: *Italics*: suggested additions
 [Square brackets: moved up the logframe]

	Indicators of Success	Monitoring & Evaluation	Assumptions
Project Goal:	Increase development effectiveness through <i>use of better evidence</i> in developing countries	MDG indicators Paris monitoring	There are substantial global public good gains to be realized by high quality, rigorous impact evaluations.
Purpose: (3ie mission)	IE Evidence influences policy and implementation	Change in policies and programs based on evidence from quality impact studies [(Obj 1&3): No. of cases in which policy or program design demonstrably affected by IE/SR findings] Increased awareness of, and demand for, rigorous impact evaluations <i>No of Enduring Questions where new thinking results from 3ie outputs.</i> [(Obj 1,3&4): No. of LMICs with on-going IEs] Awareness amongst policy makers and development practitioners of 3ie and the services it provides [The number of institutions and policy-makers reporting having been influenced by 3ie-supported evidence]	- <i>Policy makers are not too constrained by political and fiscal factors</i> [There is scope for policy to be influenced by evidence, rather than being determined solely by political considerations.] [Impact studies have the capacity to identify welfare-improving policy changes]
Objectives:	1. Promoting the generation of new evidence from IEs which is relevant to policy and program design and implementation.	No of IEs - completed (on-going) - <i>published in peer reviewed journals</i> - <i>published w/o peer review</i> - <i>Citation scores</i> <i>Windows completed: applications, grants (value)</i> (Obj 1&4): proportion of grants including LMIC researchers and real country and/or NGO engagement Satisfactory rating in grant applicant and other surveys	- Sufficient technical expertise to conduct IEs on the scale envisaged by 3ie. — Studies supported by 3ie will appeal to policy makers. - 3ie will be in a position to lever both the quantity and quality of impact evaluations in other agencies.

	<p>2. Synthesizing and disseminating existing evidence in user friendly and policy relevant formulations</p>	<p>No of SRs</p> <ul style="list-style-type: none"> - completed (on-going) - <i>published on Cochrane/Campbell</i> - <i>published in peer reviewed journals</i> - <i>published w/o peer review</i> - <i>Citation scores</i> <p><i>Windows completed: applications, grants (value)</i></p> <p>Satisfactory rating in grant applicant and other surveys</p>	<p>- As for objective 1, applied to reviews rather than primary studies</p>
	<p>3. Building a culture of using new and existing evidence in the development and implementation of policies and programs.</p>	<p>No. of studies published on 3ie website with actionable policy recommendations</p> <p><i>No of events, attendance</i></p> <ul style="list-style-type: none"> - <i>Conferences</i> - <i>Donor briefing/training</i> - <i>LMIC briefing/training</i> - <i>Seminars</i> <p>No of Institutions revising evaluation guidelines, policies or budgets on 3ie guidance</p> <p><i>Website/publication usage</i></p>	
	<p>4. Developing the capacity to promote, use and undertake impact evaluations.</p>	<p><i>No of IEs commissioned by LMIC policymakers with 3ie support (PW)</i></p> <p><i>No of LMIC initiatives directly linked to 3ie support/advocacy</i></p> <p>Percentage of proposals to OW from LMIC institutions (lead PIs); (sole PIs)</p> <p>[No. of PPGs and proportion which submit successful applications]</p> <p>(Obj 1&4): proportion of grants including LMIC researchers & country/NGO engagement</p>	<p>International researchers will be willing to enter collaborative partnerships</p> <p>Funding available for capacity building activities</p>
	<p>5. To build the capacity of 3ie and its staff to achieve their strategic objectives and improve their performance and to ensure that the systems and capacity required for the governance and management processes required by the Initiative are in place and well-functioning</p>	<p><i>3ie core funding</i></p> <p><i>Workplan objectives delivered</i></p> <p><i>Membership:</i></p> <ul style="list-style-type: none"> - <i>Bilateral/multilateral</i> - <i>LMIC Government</i> - <i>NGO</i> <p>No. of staff in place</p> <p>No undertaking staff development</p> <p><i>No. of Board actions completed/uncompleted</i></p> <p><i>Audit report</i></p>	<p>3ie will attract sufficient funding to establish itself and its credibility.</p> <p>It will be possible to attract staff to 3ie who are of sufficient calibre</p>

6.3 The Potential Impact of 3ie Grants

431. 3ie's Theory of Change, indeed its whole reason for being, rests on two pillars:

- Rigorous impact evaluations are undertaken
- Existing evidence is analysed and synthesised

432. The word 'quality' is implicit in both. To have policy impact, IEs must produce evidence which is a clear and credible addition to the existing body of knowledge. The systematic review process is designed to analyse and synthesise existing evidence while explicitly excluding evidence which does not meet stringent standards. Recent very public debates over the evaluation of the Millennium Village Programme have highlighted how any evidence on development will be 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 a body with a specific mandate for standard setting, quality and credibility are particularly important for 3ie.

433. It is too early to assess 3ie's overall impact on policy. With 7 IEs and 9 SRs completed, the evidence it has generated is only just beginning to enter the policy process. However, it is not too early for an assessment of two characteristics which, as noted above, are better determinants of the policy impact of 3ie's work than its work on policy influence: quality and policy relevance. This section presents an initial assessment on that basis. It is important to emphasise that this has been prepared by a small number of specialists, with just four days to review the complex and highly technical material behind 15 to 20 grants each. The three-person Evaluation Team has had slightly longer, but most of their time has been spent on other aspects of the process evaluation. Inevitably, the analysis reflects the personal experience and judgements of the specialists involved. For reference, the reports of the individual specialists can be made available.

434. With those qualifications, 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. In essence this section is an effort to respond to the comment from 3ie donor members' reported in the Leadership and Organizational Performance Review (see section 3), that *"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."*

6.3.1 The Meaning of Rigour

435. The Strategy 2010-213 sets out a clear statement: How 3ie Defines Rigorous Impact Evaluation. While emphasising that meeting the attribution challenge is likely to require a statistically rigorous experimental design, 3ie advocates theory-based evaluation:

"Ensuring policy relevance is facilitated through the adoption of a design that analyzes the intervention across the causal chain, from inputs to impacts, sometimes referred to as theory-based evaluation, rather than the black box impact evaluation approaches which only report mean differences in outcomes between treatment and comparison groups. Thus, 3ie supports studies that answer the question "what works?" as well as "how?" and "under what circumstances?" Unpacking the causal chain in this way requires the use of mixed methods. We also encourage the use of cost effectiveness or cost benefit analysis."

436. To support this, 3ie have produced an extremely powerful set of practice papers and guidelines underlining how rigour is not just about the statistical design. Two guidelines are recommended to grant applicants: Impact Evaluation Practice and Principles of Impact Evaluation. The Evaluation Team have particularly noted two other documents: 3ie Working Paper 6, A Checklist for the Reporting of Randomised Controlled Trials of Social and Economic Policy Interventions in

Developing Countries, and a presentation on Mixed Methods by the ED which opens with a slide “Measurement is not Evaluation”.

437. 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 be underselling it. There has been a tendency to present 3ie as an advocate of Mixed Methods, or Qualitative Methods, which are much weaker concepts. A recent paper on ‘Cause and Effect in Small-n Impact Evaluations’ may add to an impression that 3ie is only advocating mixed methods for special circumstances. The remark noted in the Rapid Institutional Review, that 3ie is not ‘walking its own talk’ on mixed methods perhaps reflects a similar feeling.
438. In the review of outputs which follows, the Evaluation Team have sought to judge grants funded by 3ie against a 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.
439. As far as is possible on the material available, the review has also looked at two other aspects. First is the quality of the grantees’ understanding of the context. Development is inherently context specific: interventions which succeed in one country fail in another, and impact evaluation designs which work in some places do not work in others. Naïve identification of confounding factors is a particular risk if the context is not well understood. The second aspect is the quality of implementation. A weak questionnaire, or poorly trained and supervised survey enumerators will undermine the strongest statistical design; or, worse, generate results which are convincing because they are statistically significant, but are fundamentally unsound.

6.3.2 Types of Intervention

440. The Evaluation Team have heard 3ie criticised as being more about ‘proof of concept’ than the evaluation of a major development programme or policy. It is certainly correct that a significant proportion of 3ie grants are tests of proposed new development interventions, or implementation approaches. This does not necessarily mean they are not policy relevant, but the distinction is an important one. A simplified classification might be as follows:
- Studies of the impact of major on-going programmes, such as Ethiopia’s Poverty Safety Net Programme, Mexico’s 70 y Mas pension programme and W. Bengal’s metering of electrically powered tubewells.
 - Studies to test the introduction of a policy which is already widely recommended in a new area or country. A 3ie example is the evaluation in Kenya of Intermittent Screening and Treatment for malaria among schoolchildren, an intervention which was already part of national policy.
 - Studies to test a relatively untried product or implementation approach. 3ie examples include cook stoves in Ghana, mobile banking in Sri Lanka and a new pension product for micro-finance customers in India.
441. It is far easier to organise a clean and statistically effective design for studies in the second and especially the third category. For the first, it is difficult if not impossible to achieve the RCT ‘gold standard’. Even quasi-experimental designs can pose problems. Conversely, immediate policy relevance is much more likely with the first category.

Implicitly or explicitly, a significant number of 3ie grants 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. Here too the advantage is that these issues are amenable to controlled-trial forms of evaluation, but it risks limiting attention to a rather short length of the causal chain.

6.3.3 Specialist Review of the 3ie Grant Portfolio

442. Table 30 presents a breakdown of the 3ie IE portfolio, by grant window, region and sector. It shows a reasonably balanced spread between Sub Saharan Africa and S. Asia, with smaller numbers in E. Asia and S. America. The majority of grants are grouped in five sectors: health, social protection, agriculture, education and private sector development, in order of importance. A number of grants overlap the sectors and 3ie shows these in more than one category. For the purposes of this review, the evaluation team have sought a classification based on the main thrust of each study.

Table 30: Breakdown of the 3ie Portfolio of Impact Evaluations

Window	IEs	Region	IEs	Sector	IEs
OW 1	17	SS Africa	36	Health	17
OW 2	30	E. Asia	8	Agriculture	12
OW 3	21	S. Asia	23	Education	11
SPTW	10	S. America	10	Finance/PSD Sector	9
		Multiple	1	Environment	5
				Governance	5
				Social Protection	14
				Other	5
				TOTAL	78

443. Triple Line has asked five independent specialists to provide an assessment of the five largest sectors in the 3ie portfolio. The specialists were specifically chosen as leading development practitioners, not as evaluation specialists. For each sector, they were asked to provide an overview of: a) enduring Issues which are most critical, or most under discussion, and which are most relevant to policy makers and donor planners; b) development models and theories of change; and c) key contextual factors. Against that framework they assessed each grant on four points:

- The validity of the theory of change
- The understanding of the context
- The potential contribution to current knowledge
- The likelihood of generating useful conclusions for policy makers

444. Systematic reviews were assessed on the same basis, but concentrating on the last two points: the contribution to current knowledge and the likelihood of policy relevance.

445. In conclusion, the specialists were asked to review the portfolio as a whole against four points:

Coherence: As a portfolio, as opposed to a collection of separate pieces of research.

Potential impact: the proportion of the portfolio which is of high quality and relevance, and likely to make a significant contribution to 3ie’s overall goal of “improving lives through impact evaluation”

Particular Strengths and weaknesses: Theory of Change; Context; Contribution; or, Relevance?

Focus: Could the work be better directed towards enduring issues and known policy questions?

446. With only four days for each specialist, and a variety of material to consider, from recently awarded SPTW applications to final reports from OW1, the specialist reports can only be impressionistic; the summary presented here even more so.

447. Table 31 presents the specialists' ratings against the four assessment points on a four point scale.³³ The scale was drawn between a low-point definition, shown at the top of each section in the table, and a high-point definition, shown at the bottom. The following summarises some key facts in the data:

- The average score across all sectors and assessment points was 2.6, i.e. somewhat on the positive side of neutral but not massively so. At 1.7, the average for Finance/PSD was particularly low, but no sector scored higher than 2.8.
- It is perhaps more useful to look at number of grants scoring 3 or 4 across all four assessment points, taking that as a measure of a truly strong IE study. 17 out of 63 (27%) grants met that standard. (Only one grant rated four 4s.) At the other end of the scale, 13 grants did not score more than 2 on any point. Of these, six were in Finance/PSD, three in Health, three in Social Protection and two in Agriculture. Although Education had no grants which failed to score at least one 3, it also had the only zero score in the whole sample, and only one top mark 4.
- Of the assessment points, ratings were lowest for a strong, testable Theory of Change. Only four grants got the top rating, all in Agriculture, although a majority of Social Protection grants scored 3. Across all sectors, 56% scored 2 or less on Theory of Change.
- For the other three points, ratings were evenly spread with the largest single group rated 3.

448. Of the six SRs reviewed, four scored 3 for contribution to knowledge and four scored 3 for policy relevance. There were only two top marks and two 2's: one for knowledge and one for policy relevance. There were no very low scores.³⁴

449. The remainder of this section presents a summary of the specialists' conclusions, sector by sector. The majority of the Social Protection portfolio was commissioned through the Social Protection Thematic Window. This reviewer's conclusions are presented separately, in section 6.4 as part of a discussion of the advantages and disadvantages of thematic windows.

³³ The scale was set at five points, but only there was only one zero rating across all grants and all assessment points.

³⁴ Not all the completed SRs were scored, because some did not fit well in the sectors reviewed.

Table 31: A Specialist Practitioner Assessment of 3ie IEs

	Fin/PSD	Health	Educ'n	Agric.	Social P	All
Score	n	n	n	n	n	n
Theory of Change: naïve based on simple intuition with no understanding of the sector						
1	2	2	1	1	0	6
2	5	8	8	4	4	29
3	3	4	3	4	10	24
4	0	0	0	4	0	4
Theory of Change: strong and testable, clearly aware of current knowledge and past experience						
Context: no attempt to take account of it in the research design						
1	5	2	0	2	0	9
2	4	1	1	6	5	17
3	0	3	10	3	6	22
4	1	8	1	2	3	15
Context: research which is fully tailored to the specific context						
Knowledge: tests issues that are already well understood or of little or no current importance						
1	8	2	0	0	1	11
2	1	2	3	4	3	13
3	0	3	9	6	10	28
4	1	7	0	3	0	11
Knowledge: directly addresses at least one enduring issue in the sector; if successful, will substantially increase understanding of that issue						
Relevance: will only generate results of academic interest, if at all						
1	6	2	0	2	1	11
2	2	2	6	1	4	15
3	1	3	6	4	8	22
4	1	6	0	6	1	14
Relevance: if successful, will allow policy makers to substantially improve policy or implementation						

Private Sector Development

450. The reviewer did not find significant coherence in the Finance/PSD portfolio. There was little linkage between the projects beyond the use of randomised controlled trials throughout. Designs are high quality but the theory of change is never explicitly stated and this has led to a general lack of clarity as to the purpose of the project and how its findings are expected to inform policy makers usefully. The emphasis seems to be on results of academic interest, rather than on practical and usable guidance to those designing private sector or financial development initiatives.
451. Contextually there is significant naivety: it is frankly unrealistic to expect informal sector microenterprises to maintain accurate records of business and household income and expenditure over extended periods. The unreliability of that data is likely to call into question the carefully designed edifice of the randomized controlled trials. In several instances, the main beneficiary seems to be a single private sector bank or mobile phone company. This might be justified on if it is pro-poor and replicable, but there does not appear to be such scope for replicability.
452. The portfolio is particularly disappointing with respect to private sector development with almost no consideration of the major discussion on “making markets work for the poor” – probably the most debated area at present and certainly the area with most significance to donors in this field.
453. The PSD specialist looked at 3ie’s own reviews of the submitted proposals. He found these technically excellent, but lacking the specific sectoral input which might have highlighted some of the contextual and relevance weaknesses.

Health

454. The Health reviewer also found her sector portfolio to be a collection of separate research studies rather than a coherent research portfolio. About a third of the grants seemed to be of high quality and relevance (especially the Matlab study on long term impact in Bangladesh, TB patient compliance in Pakistan, health impact of sanitation in India). Another third are high quality studies but on topics which are a small element of a sub-topic, nothing approaching an enduring issue. Although some studies have engaged country stakeholders and policy makers, most seem to be driven by outside researchers (mostly US universities) often with only a token national team member. The three strongest proposals are evaluations of on-going and well-established programmes, not new interventions. All of them seemed to have a good understanding of the issue and country context, but several assumed their findings would be generalizable elsewhere.
455. There is a disproportionate emphasis on fortification of foods (6/14); and Theories of Change are unexplained, beyond brief references to causal links. The reviewer noted that a limited number of organisations are involved in the portfolio, with some holding several different projects. She was particularly struck by the fact that only one of the grant holders or researchers was a School of Public Health. Instead most are Departments of Economics. Several proposals involve health education (for nutrition or family planning), but the world leader in this area, the John Hopkins Bloomberg School of Public Health, is not involved.
456. The reviewer thought the portfolio would be strengthened by focussing on specific health themes and issues, perhaps linked to donor priorities for health research. It is often said that evidence exists about cost-effective interventions but the problem is how to apply them. Potential focus issues include scaling up proven interventions countrywide, and health systems strengthening.

Education

457. Although the Education portfolio covers primary, secondary, VET and higher level education, ECD is over-represented. Studies focus on access to education, with limited attention paid to survival and retention and to different dimensions of quality. Equity issues are only addressed through scholarships or CCT for the poor, that is through a direct input to the end beneficiary. Other approaches such as better teacher deployment in remote and poor areas, or pro-poor school grant schemes are not studied. With three exceptions, the evaluations focus on similar interventions (Cash/CCT, scholarships or day care), without attempting to evaluate alternatives, or provide a comparison of policy options. Other aspects of education, textbooks, teachers, school management etc., are not addressed. The fact that five of the OW grants are in four of the nine most heavily populated countries is a clear strength. Many of the evaluations look only at short term impacts. While this is understandable because of funding constraints, education cycles run over several years and outcomes achieved at a young age might not significantly affect students' schooling over time.
458. The OW and SR portfolios are quite complementary, and the SRs help to highlight geographical gaps, intervention characteristics gaps and methodological gaps. It would be useful to take these into account when OW proposals are assessed. This would increase the quality and comparability of the evaluation outputs.
459. Many proposals indicate that qualitative data on interventions will be collected, but how this mixed-method data will be used to complement the quantitative findings is unclear. There is little evidence of this being done in the draft/final reports completed so far. The majority of the proposals are not framed in terms of a Theory of Change. Links between intermediate and final outcomes are not systematically explored, nor is the effect of contextual factors. Most applicants are familiar with the country context, having worked in the country before and secured partnerships with government or key agencies. Despite this few proposals present a clear summary of the policy context, and questions of political will, the role of the media and other social factors tend not to be taken into account.
460. How proposals address cost-effectiveness and cost-benefit varies. There does not seem to be a clear definition of the concepts and of a standard methodology. From the two final reports available, cost-effectiveness is an area in which there is a discrepancy between what the grant proposal promises and what is actually delivered.
461. In summary, the Education IEs address the "What works?" question. Some also look at "For whom?" by disaggregating data by gender, location, poverty levels, etc. They address the questions of "Why?" and "Under what circumstances?" far less effectively. There is, for example, limited focus on how intervention processes affect delivery and the results. The reviewer feels, therefore, that the 3ie outputs may well contribute to knowledge creation but some of them may offer limited policy options. Although some evaluations assess the health and economic returns of an investment in education, this is only useful where policy makers need to be convinced to make that investment. Most countries have education sector plans that cover all education sub-sectors. What policy makers need most, is policy options and alternative intervention models with clear analysis of the cost effectiveness, to support their decisions on how best to invest each dollar. As a final note, the reviewer felt that the high academic drive of the evaluations has resulted in reports that can only be accessible to highly educated donors and policy makers.

Agriculture

462. The Agriculture sector reviewer identified three IE studies of combined high quality and relevance, six in a middle group and four below average. She noted that two of the lower scores were in agricultural extension. Although six of thirteen were ranked as 'if successful, will allow policy makers to substantively improve either policy or implementation', half of these showed

limited contextual understanding, which may limit sustainability, and weaken links between outcomes (e.g. credit uptake) and the impact on livelihoods. Most differentiate beneficiary groups by household type and farm size and some within the household. However, 3ie has not registered any as having an explicit equity focus, and only five as having a gender focus. Given how central equity and gender are to the international agenda on rural development, the reviewer found this surprising.

463. Most of the agriculture studies are linked to existing policy changes or programmes. This gives them the potential for real policy impact, at least in the country concerned and especially on some larger programmes: The World Bank's \$163 million Bihar Rural Livelihoods Project; Farmer Field Schools in China; and, Metering of agricultural tube wells West Bengal. There is growing international interest in new models for weather and market risk management. Four IEs related to these topics have the potential to contribute to the national and wider debate.
464. Among the smaller number of less relevant IEs, the reviewed noted one on Agricultural Technology Adoption in Malawi as an example of the weaknesses. It looks at how innovative delivery mechanisms affect the uptake of two conservation agriculture practices. Whilst no doubt of academic interest, the IE fails to explain how it might specifically contribute to current range of delivery mechanisms. It is an expensive study, and it would have been more useful to analyse the effectiveness and efficiency of Malawi's on-going extension models.
465. The reviewer reports no explicit presentations of a Theory of Change, although three IEs make the causal linkages clear within an explicit testable theory. The ToC implicit in the five weakest are based on no more than intuition, with outcomes that are not directly testable against the intervention under study. One study, on Low Cost Farm Equipment in Africa, is highlighted for the simplistic nature of the link between intervention and outcome and the weak articulation of the sector.
466. Nine of the Agriculture IEs are RCT studies, four of them with some mixed/qualitative methods as well. This may answer the 'what works' question, but it will not reveal why have some farmers benefitted and others did not – what are the preconditions and barriers? The short timescale for the trials, often 12-18 months, and in some cases the intervention under evaluation raises questions over sustainability of the impacts/outcomes measured. It also makes some proposed outcome measures unrealistic. (Again the Low Cost Farm Equipment study is noted as an example. With an early buyer cohort in 2010 and a late buyer cohort in 2012, it seeks to measure impacts on livelihoods, education and health. The study team note this may be unrealistic, but it remains embedded in the overall aim of the IE.) The only exceptions to this are one post-intervention IE on irrigation improvement and one three-year study of an agricultural voucher scheme in Tanzania.
467. In summary, the IE focus on specific, limited interventions and the measurement of outcomes over a short period fails to recognise the complexity of the sector and the multiple needs of rural households and small-scale farmers. This risks raising the profile of specific interventions to the exclusion of a wider understanding. 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 of the intervention.
468. The Agriculture sector reviewer makes a number of short points which are worth reviewing, partly to demonstrate how 3i addresses many of them, as detailed in the footnotes.
469. **Transparency and access** - Publishing IE protocols on the 3ie website would enable peer comment, raise interest and help to show that the outcome reported was the outcome proposed.³⁵

³⁵ The Evaluation Team believe that grant proposals were published on the old 3ie website. Clearly this suggestion links to 3ie's proposed IE Registry.

Conflict of Interest – The reviewer notes a number of IEs with members of the research team who are from the agency implementing the intervention. A particular issue concerns people who are developers or sellers of a commercial product under evaluation. While this may raise ownership of the study, it must raise questions as to independence and issues over product endorsement, commercial confidentiality and open data access. If allowed, this involvement should be made explicit in the protocol and reports.³⁶

Timeliness of outputs – A number of IEs require tight time delivery of outputs to meet policy or investment decision points. This should be noted by 3ie in contract negotiation and in monitoring that the research teams deliver within agreed timetables.

Capacity building - The reviewer suggests 3ie should score this aspect of IE applications much harder, and require a budget for capacity building for national researchers. The current proportion of national researchers is too low.³⁷

Language – The reviewer notes a strong bias towards English speaking researchers and countries, with francophone W. Africa a particular gap.³⁸ It would also be useful to bring in country experiences from, for example, Latin America.

Ethics of RCTs - For non-treatment cohorts there remain equity issues on terms of access to the services offered. Gifts to part of sample households, within a treatment and within the same community, (OW3.1243) must raise ethical questions.³⁹

The extractive nature of evaluation - Data collection is time consuming and for non-beneficiaries (assuming benefit of intervention), they stand to gain nothing but loss of their time. The Evaluation Team have noted grants with extremely large questionnaires, (76 pages – OW1.43), frequent repeat surveys, and one where blood samples were taken without, judging by the wording of the questionnaire, clearly seeking the interviewee’s permission (OW1.43). This concern deserves attention.

6.3.4 Statistical Review of Completed IEs

470. 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/draft final reports of nine 3ie studies. After confirming that 3ie’s guidelines provide clear procedures for planning and executing rigorous IEs, the specialist undertook a detailed review on all aspects of the reports including quality of the presentation, rigour of the design, methods of data collection, statistical analyses and the relationship between the results and the conclusions. 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?

471. The results were as follows:

1. Paying for Performance in China’s Battle Against Anemia (OW 1.66)

The policy recommendations seem unjustified given the results, although the lack of robust analyses may mask real differences between arms.

³⁶ The Finance/PDS reviewer raised similar concerns.

³⁷ 3ie strengthened its requirements in this area between OW1 and OW3/SPTW

³⁸ The Education review also noted French W. Africa as a gap

³⁹ 3ie specifically requires grantees to provide evidence that their proposal has been reviewed by a qualified ethics board. 3ie reviewers raised this concern on a number of grants, but ethics board approval seems to be taken as adequate.

2. A rapid assessment RCT of improved cooking stoves in rural Ghana (OW 1.69)

This report is poor – the design, sample size, methods of evaluation and statistical analyses all lack rigour.

3. Does Marginal cost pricing of electricity affect groundwater pumping? (OW2.082)

The results presented here are very difficult to follow and most of what was presented is not significant. It does not appear that a power test was undertaken to work out what sample sizes would be needed to show a significant difference. There is a need for a much clearer explanation of how the data were collected and analysed. All the information about benchmarks appears not to have been used. It remains unclear whether or not farmers are changing their groundwater pumping behaviour as a result of the introduction of metering.

4. Impact of malaria control and enhanced literacy instruction on educational outcomes among Kenyan school children. (OW2.065)

Normally a report would present Findings – Discussion – Recommendations. But this one has no Discussion and so it does seem a rather sudden leap to a recommendation not to implement IST in schools in Kenya (p27). It is also confusing why they make this policy recommendation now, after 12 months, when they obviously continued implementing the IST for malaria intervention for a further 12 months (“the 24 months follow-up results are pending”, p27).

Tables 2 and 3 do not appear to correct for age which is known to impact on malarial prevalence. All the analyses presented here are very simple but as we do not know if there are significant differences between intervention groups in for example age, educational level of household etc. then it is impossible to judge whether more sophisticated analyses should have been undertaken.

This is an interim report covering the first year only and the full impact of the study will not be known until after the 24 month follow-up. Much more detailed statistical analyses need to be undertaken before one can judge the effectiveness of this programme.

5. Evaluating Indonesia’s unconditional cash transfer program 2005-6 (OW 1.76)

A well written and sound report.

6. The promise of preschool in Africa: a randomized IE of early childhood development in rural Mozambique (OW 3. 1229)

This report leaves more questions than answers! No rationale for undertaking quite complicated statistical analyses is given, when easier before/after or differences-in-differences models seem more appropriate.

7. The impact of day care on maternal labour supply and child development in Mexico (OW1.43)

This report is unnecessarily detailed and it could be pruned by about a third and still maintain the salient components. There is potential to produce some succinct research papers.

8. Program 70 y mas (OW3.1229)

The conclusions are dubious. Even if analysed correctly, it is questionable whether data could provide convincing evidence for universalising the programme.

9. No margin, No mission (OW 1.7)

Although at first sight the results are counterintuitive, the specifics of the intervention, the financial rewards and the low sales need to be taken into account when drawing conclusions.

472. The assessment above would suggest that only two of the nine papers are definitely publishable. This assessment is disappointing. 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. (OW1.66) 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."⁴⁰
473. Nevertheless, this 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.

6.3.5 IE Implementation Review

474. The evaluation team has 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.⁴¹ 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, (for completed grants). This was to identify whether things ran smoothly, or whether there had been minor or major issues with the grant at each stage.
- i) The proposal and application almost by definition went well for the selected grants. However, a third of the grants looked at had minor issues. Some OW1 applications had to be revised, one to delay the grant by up to a year. Overall grantees dealt well with the 3ie application processes, including the re-submission of revised applications. Four grants were substantially criticised by one or other of the external reviewers for lack of technical rigour, or relevance, including one which the policy reviewer had recommended as a reject.
 - ii) The Conditions and Approval process included negotiation on the C4r, BC4R, and the PIP for OW3 grants. Cost was an issue on four of the grants, with 3ie-GDN seeking reductions or more detailed explanation in many cases. In a few cases there were several rounds of comments for response and grantee feedback, causing some delays. To sum up, there were minor issues on 2/3 of the grants, some involving extensive comments for response and the grantee not responding satisfactorily. In two cases the 3ie EPA was critical of other reviewer's comments.
 - iii) Contracting was an issue in a quarter of the grants looked at, and half requested at least one letter of variation. As mentioned, US institutions had problems with some bits of the contract, so had to transfer the grant to IPA. This was a major source of delay.
 - iv) All but one of the grants had minor issues over Deliverables and Disbursements. Three of them had major issues involving: extended delays, which almost led to grant cancellation; late reporting of problems with survey implementation; and repeated inadequate responses to 3ie comments and suggestions.
 - v) Of 10 completed grants, four had no issues. Of the others, two draft reports were delayed and four were heavily criticised by 3ie's external reviewers. One concluded that all that could be learned from the project is that "badly designed and implemented studies don't tell us much of use."

⁴⁰ The full review can be found in Annex 2.

⁴¹ The Evaluation Team asked for the records for 21 grants. Only 20 were supplied and the record for one (OW1.45) was without an audit trail of comments, responses and approvals by 3ie ER and EPA.

475. Overall, 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, that so many of the sample show problems at this level is cause for concern.
476. To try to throw more light on this, the evaluation team leader personally reviewed 11 grants: eight from the OWs 1-3 and three from SPTW. In the rest of this section, he presents a summary response to the OW grants. (For the SPTW he has nothing to add to the sector specialist's report.) Given the time available and the volume of material, these responses can only give an impression, and they are subject to checking. They are snap judgements by an experienced practitioner and should be treated as such.
477. **OW1.7 – No Margin, No Mission.** This completed study aimed to test the impact of cash and social incentives on the performance of community agents, a very important question in Community Driven Development. In practice, it measured how Zambian hairdressers, not community agents in any real sense, responded to incentives for the sale of female condoms. 3ie was co-funder. The survey was carried out by the implementing agent, under contract to 3ie, with the PI contracted separately through IPA. On completion the grantee submitted a nine-page final report together with a set of dissemination materials in December 2011. 3ie informed the grantee this was not enough to allow the technical quality of the work to be confirmed. It was then discovered that without informing 3ie the PI had been circulating a working paper on the internet for some months. At 3ie's suggestion, this working paper was accepted as the final report. The short published paper claims to have shown that social incentives are more powerful than cash incentives, but it makes no mention of the statistical significance of the result. Perhaps more importantly, it admits that even the best performing hairdresser only sold one pack of condoms a month, rendering the result largely irrelevant. 3ie's reviewers had highlighted the need to look at 'present up-take rates', and the fact that delivery is unlikely to be the critical determinant of up-take. In hindsight, those criticisms should have been taken much more seriously.⁴²
478. **OW1.43 – Daycare, Maternal Labour and Child Development.** This grant was to fund a follow up survey for a baseline study of a large Mexican government daycare programme which was already reaching 50,000 mothers. The survey design assumed excess demand for daycare places, so that a control group would be found among those who were not randomly assigned a place. This assumption failed drastically, with only 648 applications instead of 4,500. An apparently acceptable way was found to get round this statistically. However, the stated assumption that "households not interested in the program are not a good comparison group, given the number of variables of interest in which there is no balance between the groups" is revealing of the way statistical issues seem to have obscured the policy-relevant elephant in this particular room: the lack of demand. Even with the new approach, the survey sample included 16% replacements, and a further 22% replaced through attrition: further indications of weak demand. This is the first of a number of studies where the PIs' reconnaissance was inadequate to spot major issues. A 76 page questionnaire, including blood tests, risked being a real imposition on those interviewed.
479. **OW2.161 – Economics and Psychology of Long-term Savings and Pensions.** This study seeks to measure how different designs of pension product affect uptake among 6,000 savers with a large Indian NGO: 'The product design effect on savings participation.' It is expected to answer questions on time preference, risk aversion and psychological barriers to savings. It involves a baseline survey and seven follow-ups at 3-4 monthly intervals. With a start date of April 2010, the first progress report was not required until February 2012. The grantee submitted a comprehensive set of questionnaires and has completed the baseline survey. However, the

⁴² The Evaluation Team have been unable to review the final version of the Working Paper because it is behind a pay wall.

intervention was stalled because NGO sales staff would not promote the pension product without a commission. Despite promising this, the NGO was not willing to pay it. This seems odd. The grant application to 3ie stated that the NGO was already offering pension products to poor women. No details of the pension product were given in the application, or in the progress report. As with other grants, a statistical work-around was proposed but the EPA pointed out that it was not clear it was effective, and what impact it would have on power calculations etc. Nevertheless, he approved the 2nd tranche payment, saying the set of deliverables was fine. These included a 66 page questionnaire, with a minimum time set for application of 2 hours 35 minutes. Many of the questions seem unworkable. “What will you spend the most on in the next 5-10 years? - Multiple answers allowed,” is logically incoherent; and almost impossible to answer from 14 options ranging from house purchase to daily consumption over 5-10 years. Even less convincing is the attempt to apply behavioural economics tests on the preferred cash outcome from a coin toss to poor Indian women. The conclusions are exactly the same as for OW1.43: poor reconnaissance and an impractical questionnaire. The lack of reconnaissance was confirmed by an early report that “individuals were more exposed to financial products than assumed.”

480. **OW2.065 – Malaria Control and Enhanced Literacy Instruction.** This study was a full-scale RCT testing the impact and interaction of intermittent screening and treatment (IST) for malaria and enhanced literacy instruction across 100 Kenyan schools. The proposal was strong and the statistical design was delivered as proposed. Although some 3ie reviewers seem to think the study is finished and a draft final report has been produced, a second follow-up survey for literacy has yet to be completed. The results so far show no statistically significant IST impact on malaria infection or anaemia rates. The education intervention did show an impact on Swahili literacy. There was no interaction. The report includes a cost-benefit analysis and a perception survey to test public reaction to IST, claiming this as a mixed-methods approach. It recommends that IST should not be implemented in Kenya schools. 3ie reviewers suggested the evidence was not strong enough for such a recommendation, in particular because there was a possibility of poor compliance with the three-day course of malaria treatment. While this is one of the strongest studies produced so far, the compliance question highlights a critical weakness in the causal chain, one which might have been expected. There are other questions which might have been followed up, in particular the fact that malaria infection rates varied between zero in seven schools and over 40% in three. It would be interesting to know how this heterogeneity affected the statistical power of the study, especially so since the impact on malaria infection rates was large enough to be policy relevant, a 25% reduction, even though it was not statistically significant.
481. **OW2.082 – Electricity metering of agricultural tubewells.** This study sought to measure how W. Bengal’s replacement of flat-rate electricity tariffs with metered rates had affected water markets. The final report was published in November 2011. It reported that the “expected impact on reducing pumping hours was felt only in the boro season,” and “There is also some evidence ... that water sales and purchases were also adversely affected.” There was no change in cropping patterns or boro paddy output, a fact explained by the suggestion that farmers were overwatering prior to the introduction of metering. This reviewer has reservations about this study:
- Metering is designed to give the electricity company the incentive and the resources to improve supply. The report highlights how powerful supply constraints had led to rationing so that poor quality electricity had ‘become the norm’, with hours of supply reduced by as much as two-thirds in some states between the 1980s and the 2000s. Without thorough discussion of the supply side of the electricity market, it is not enough to suggest that difference-in-difference can ‘wash out’ the effect of supply constraints.
 - The grant application offered a natural experiment based on pre-metering baseline surveys and the phased roll-out of metering. The survey as delivered was more complicated. The

study covered two ecological zones – old alluvium and new alluvium. Metering coverage was higher in the second, with the result that the metered/unmetered sample was unbalanced across the zones. To control for this the sample was ‘augmented’ with a further 18 villages. In the Gangetic plain zones which by their names might seem very similar can hide differences in soils and ecology which have a large effect on agricultural technique, in particular on crop water requirements. This is a large confounding issue which should have been discussed in more detail.

- Among the large number of relations tested, the single statistically significant result related to boro hours pumped by the owners of ‘either electric submersible or electric centrifugal pump sets’. The difference between a centrifugal and a submersible pump is very large indeed. The first can only lift water 25 feet and irrigate perhaps five acres. On a deep tubewell, the second can lift from hundreds of feet and irrigate up to 100 acres. Any statistic which blends the pumping hours of both classes of pumps will almost certainly be dominated by the mix of pumps, not the hours pumped. That the result should be almost equal, yet statistically insignificant, for submersible pumps alone seems quite unlikely.
- The likely impact of metering is modelled as a matter of very narrow, yet highly abstract behavioural economics: essentially as a two-person game between the pump owner and the farmer. This ignores 20 years’ worth of more straightforward micro-economics which has explained the development of Bengal and Bangladesh water markets very convincingly. In one sentence, rapid expansion of a highly competitive pumping industry has driven marginal returns to water down to zero, so that factor returns to land, labour and water are balanced. If there is an interesting behavioural question, it concerns the disappearance of water share-cropping, which has clear issues over asymmetric knowledge, and its replacement by a flat-rate seasonal charge. Even the report’s proposed monitoring assumption, that water sellers have a problem monitoring water buyers’ use, seems strange. A meter may solve that, quite apart from the unsupported assumption that pump owners just hand over the keys of their pump to the farmer.
- Competition in the water market is shaped by exit and entry. It is possible, even probable, that the largest impact of a higher electricity price was through the shutdown of marginal pumps. The study treats attrition as a problem to be handled by through statistical technique and replacement, where it may actually be the major impact outcome.
- Ignoring behavioural economics, any changes in the water market are going to be articulated through the prices of electricity and water. The report makes no attempt to estimate how the new metered rates will affect the price: a back of the envelope job given an estimate of the hours pumped under the fixed tariff.
- There is no analysis of what lies behind large secular changes in production and pumping, in particular changes in the relative prices of rice, labour, land and fertiliser. The attempt to abstract from these powerful forces using statistical technique seems unhelpful, even if it succeeds. The lack of interest in the real world is underlined by a failure to indicate how the size of the estimated reduction in pumping due to metering compares to that secular change.
- The last point is the most important of all. Table 32 below shows a small sample of the data delivered to 3ie. It shows how the hours pumped figure, is in fact, calculated from the owner’s recall estimate of days operational and hours per day. Both figures are, at best, rough estimates with only a tenuous relationship to the actual number of days and the precise average per day over a long season. The point is underlined by the fact that six of seven pump-owners reported exactly the same numbers for 2009 and 2010. One did not give an answer for 2010, presumably because he did not operate his pump at all. Unless this reviewer has misunderstood something fundamental, it is difficult not to conclude that the apparent precision of the statistical results is spurious.

482. 3ie’s peer reviewer mentioned several of these points: prices, secular trends, and that the “authors are a bit cavalier about selection bias.” But he introduces his short report by saying that his comments “are only minor.”

Table 32: Data from OW2.082 on Hours of Water pumped from Tube Wells in W. Bengal.

Q32_2Summer 09	Q32_3Summer 09	Q32_4Summer 09	Q31_2Summer 10	Q31_3Summer 10	Q31_4Summer 10
32.2	32.3	32.4	31.2	31.3	31.4
Number of days in the season when operational	Number of hours in the day when operational	Total number of hours (2X3)	Number of days in the season when operational	Number of hours in the day when operational	Total number of hours (2X3)
110	21	2310	110	21	2310
90	22	1980	90	22	1980
40	4	160	60	3	180
100	22	2200	100	22	2200
30	2	60	30	2	60
90	5	450	0	0	0
30	5	150	30	5	150

483. **OW2.192 – Micro-savings through Mobile Phones.** This study is a proof of concept pilot of a mobile phone product for a Sri Lanka bank. 3ie is providing top-up funding to extend a planned survey from 1,000 to 2,500 households. This would provide the statistical power to add “price elasticity, commitment savings, and other options to the savings menu.” This expansion was linked to a major shift in project design, from manual to mobile-phone banking. Initially, there was no MoU with the Sri Lankan partners and the mobile phone product was still under development, apparently by the research team. By the end of 2011, the main design issues had been resolved and “we began rolling out the product in the Kandy area.” This study raises two issues. Whether or not it is appropriate for 3ie to support, even if only through survey costs, the development of a potentially highly profitable commercial product, and secondly, whether it would not be more appropriate to insist that the product is fully developed and tested as a pilot, before involving large numbers in a wholly untried intervention, so they can be surveyed for experimental purposes.

484. **OW3.1118 – Seeing is Believing.** This study seeks to find out “what activities and incentive mechanisms can make [agricultural] extension services work”. Linked to a large World Bank project in Mozambique, there were to be baseline, midline and end line surveys across 200 communities and 4,000 households. 3ie is providing \$500,000 in a total \$1.5 million study. The other large tranche will come from the project. The intervention is to promote Sustainable Land Management (SLM) through demonstration plots. The experimental design compares plots run by run by community facilitators, plots run by facilitators who get a small incentive, and plots run by women. This study is, on the material available, blind to the extensive literature on agricultural innovation. Unproven assumptions include: “the size of landholding adequately proxies for farmers’ sophistication”; the most unlikely statement that “[agricultural] adoption itself has no objective monetary value”; and, “adoption is most likely a one-off decision”. The insistence that there is no monetary value presumably justifies a failure to present any evidence

on the financial costs and benefits a farmer can expect from using SLM techniques. The 3ie EPA did ask for more detail on the techniques proposed, but the reply was entirely general, unsupported by any evidence from trials or adaptive research in Mozambique. The research team reports positively, but it seems to be facing major difficulties: failure of the first survey contractor to deliver; the impossibility, for unclear reasons, of community facilitators working with 15 farmers; and significant problems with the public extension service. Changes to the study design include a reduction from three surveys to two, and reduction of the performance required of the community facilitator from 15 farmers with demonstration plots to the establishment of one plot, his own. How this will not undermine the whole design is not explained. Once again the conclusion is that reconnaissance was extremely poor.

485. **OW3.1229 – Non Contributory Social Pension Programme.** This study is to re-analyse an existing survey of 5,000 households to evaluate the impact of the pension programme 70 y más on mental health and nutrition. A draft final report has been delivered with a recommendation that “the Mexican government should invest efforts towards universalising the program 700 y más”. 3ie have sent a detailed set of comments on the draft, raising a number of issues. This reviewer concurs with the 3ie comments, noting in particular the way the researchers have imputed values for 20% of the sample, without clearly explaining the basis for the imputation, and the way difference in difference techniques are used across two samples, which are not homogenous, and only two periods. 3ie have sought to suggest solutions to these problems, but there is a risk of becoming involved in a prolonged technical exchange over a study that is ultimately unproductive. To quote the Evaluation Team’s statistical adviser, “It is unlikely that even if properly analysed too much will come out of this.”

6.3.6 3ie’s Systematic Reviews

486. The Evaluation Team have not given as much time to the Systematic Reviews as the Impact Evaluation Grants. 3ie London provided a tracking spreadsheet covering nine SR1 and SR2 grants, which gives a helpful starting point. On five of the nine, 3ie staff from the PO and SRO had to make additional efforts to progress the grant. Progress-chasing of delayed deliverables was the most frequent requirement with, for example, grants due in mid-2010 only completed in early 2012. There were two grants with more serious issues. On one SR2 study, the delay in receiving the study protocol reached the point that 3ie suggested the grant might be cancelled. On the other, technical weaknesses led 3ie to bring in new researchers to carry out the quantitative side of the work, leaving the original grantee to deal with the qualitative aspects. The first of these had an original due date in mid-2011, but the study protocol was still under review in March 2012. The second, which was due to complete in October 2009, revised to October 2011, has still not been completed.
487. Overall, 3ie has managed the SR grants robustly and this is reflected in 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 two SRs completed internally, one on water and sanitation and one on farmer field schools. However, the SR teams have needed considerable support from 3ie, including seconding 3ie staff to the study team.
488. Almost all of these are technically strong reports, applying rigorous selection to ensure that only experimental/quasi experimental studies which controlled for endogeneity and self-selection bias were reviewed. Despite searching very widely and considering up to 20,000 studies for inclusion, relatively few were found which met those standards: six for early child development; 12 for HIV prevention; 7 for social cohesion; etc. 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 way it was implemented. To take one of the larger samples, among 31 water treatment studies there were three different techniques being offered: chlorination, filtration, and ultra violet, with added complexity from with/without storage combinations.

489. To sum up, the availability and quality of source material – high-standard studies focussed on a clearly defined treatment/intervention – is one of the biggest constraints on the policy impact of Systematic Reviews. In a number of sectors, there is a risk of reaching the weak policy conclusion that more research is needed. The last paragraphs from the summaries of four out of six SRs published in 3ie livery, effectively present that conclusion: e.g., “Future willingness to pay [for cleaner water] studies should be based on real purchases and use;” “More studies on interventions to promote social cohesion are needed”; etc.
490. 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.
491. This is another area in which the Evaluation Team believes that 3ie’s message about Theory-Based Evaluation and mixed-methods is fundamental. The current systematic reviews do present a theory of change, but most are expressed in simple terms, using line diagrams and a relatively short discussion of the critical assumptions and key contextual factors. Even in this limited form they are an important part of the SR. The Social Cohesion SR, for example, does a neat analysis of how intervention characteristics may affect project outcomes. One of those characteristics is the extent to which project staff are advising, or directing what the community does. Of seven studies, the SR shows that this key factor is not reported in three, and is significant in four. It goes on to argue in the Summary that the “interventions have not been carried out in accordance with the theory [CDD interventions].” This qualitative conclusion is, in fact, both policy relevant and more important, than the finding that the statistical evidence is weak and more studies are needed.
492. That said, it is important to emphasise the important negative function of systematic reviewing: challenging Impact Evaluation practitioners to improve the quality of their work. The current debate over de-worming between the Cochrane Collaboration and J-PAL has already been mentioned. In the Evaluation Team’s view this is a good example of this process at work. In other words, 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 findings of the previous six sub-sections of this report.

6.4 Enduring Questions and Themes

493. 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. This section briefly reviews the issues and then looks at 3ie’s Social Protection Thematic Window as a case study; one which can be compared with the review of Open Window grants presented in the previous section.

6.4.1 An Enduring Question for 3ie

494. 3WEL put “Identify enduring questions about how to improve social and economic development programmes” as the first of eight specific activities for an international initiative on impact evaluation. This was carried over into DFID’s Project Memorandum committing its support to 3ie: specifically to the “implementation of clusters of new impact evaluations that systematically address questions of enduring importance to policy makers in developing countries.” The original MoU with BMGF included a similar commitment to focus grant rounds on “themes and questions.”

495. 3ie made a significant effort to meet these expectations, culminating in an Enduring Questions Workshop held in Kuwait in early 2009. This considered responses from stakeholders identifying over 700 questions, priorities identified in PRSPs and participatory poverty assessments, an online survey with 1,500 responses, and submissions from AusAid, Oxfam, and the governments of Uganda and Mexico among others.
496. The workshop identified child welfare, risks/vulnerability, governance and employment as possible themes and a range of sector/theme specific questions across: education, health, agriculture, governance, social protection, and employment. These were to *“be taken into account by the 3ie Board in determining the priority themes and EQs for 3ie’s grant program.”* In the event it is understood that the Board could not reach a consensus on the need for an approach focussed on particular sectors or themes. Some felt that researchers were best placed to identify priorities and that the more responsive approach of Open Windows, would provide sufficient strategic focus, provided that the selection took account of the policy relevance and potential policy impact of the studies proposed. There was also, perhaps, a tension between a focus on specific sectors, and a wish to look at more over-arching enduring questions.
497. This hesitation over the adoption of a strategic approach to grant selection led to some tension with 3ie’s principal funders, given that their commitments had been made on the expectation that enduring questions and themes would be directly addressed. With the possible exception of Early Child Development, the sector specialist reviews presented above 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.

6.4.2 The Social Protection Portfolio

498. Reflecting the continued debate, the Evaluation Team has been asked to review the advantages and disadvantages of the open and thematic windows, and to assess the Social Protection Thematic Window (SPTW) separately. This first thematic window was launched in October 2011. It is not being financed out of core funding, but as a dedicated window funded separately. The specialist reviewer who looked SPTW was, however, asked to work within the same framework as the other sector reviewers, but to take account of the more strategic approach. The portfolio he looked at included four grants awarded under the open windows and 10 from SPTW.
499. SPTW was managed by 3ie working closely with DFID and other donors. An Expert Advisory Group was formed with representatives from the social welfare/development or evaluation units of four donors: IADB, World Bank, DFID (3) and NORAD. The process started with a Scoping Report for the Design of 3ie’s Social Protection Window. The report was based on an inventory of social protection programmes in design or under implementation with support from the World Bank, DFID; and consultation with social protection specialists, policy makers, evaluators and academics. These were used to develop a comprehensive mapping of knowledge gaps, research questions and areas of interest.
500. A thematic window must start from an agreed definition of the theme. Definitions of social protection can be very broad, encompassing virtually all development activities, or narrowly proscribed, with a focus on social transfers. There are various frameworks defining the range of social protection measures, such as the World Bank’s Social Risk Management (SRM) or its newer *“Social Protection and Labor Strategy for 2012-2022”*. In the end, the 3ie scoping study suggests that SPTW should *“include the evaluation of unconditional and conditional cash transfers”* with specific mention of public works; suggesting that it *“may also consider complimentary (sic) interventions in the area of risk prevention (unemployment benefits, pensions and insurance) and opportunity promotion initiatives, like job creation and training”*. In essence, the focus of the thematic window should be on social transfer programmes, looking at enduring social protection issues under two headings: impacts and design.

501. Research questions about impacts include the contribution of social transfers to economic growth, purchasing power and labour market effects. Are social transfers spent locally? Do they support local markets? Do they make their recipients lazy? Or do they give them the confidence to take risks, diversify their livelihoods, seek employment, or invest in higher-return enterprises? There is also need for a better understanding of how social transfers affect education outcomes: not just attendance, but also learning. Interactions with health and nutrition are also important, and poorly understood. Another important area is the relationship between formal social protection systems and traditional support practices, in particular the gender implications.
502. In the area of design, there are key debates around targeting, the use of conditions, the concept of graduation, and how to link social transfers to complementary interventions. Although presented as being technical discussions, they are intimately entwined with the differing ideologies of development partners, and with the developing country context.
503. As to ideology, there are two main models. The first is poverty-targeted. Scarce resources will be better spent if they are focussed on the poor. Typically beneficiaries are expected to do something in exchange for the benefit. This approach relies on careful targeting, to select only the poorest, and favours the imposition of a condition on participants: either work through a public works or employment guarantee scheme, or by a set of behaviours, such as visiting a clinic or sending a child to school. Finally, this approach emphasises graduation, encouraging people to move out of the programme when they no longer need it. The second is a more universal and inclusive approach. It places greater emphasis on tackling inequality and aims to support broader groups. Benefits are entitlements for all in the group, even those that are not poor. This is inevitably more expensive. But its proponents argue that this universal approach is more popular and has much stronger political appeal. By attracting a larger share of funding, it will be more sustainable fiscally. This may mean that the poor get better, more sustainable transfer than they would from a programme targeted exclusively at them. As entitlements, transfers under this approach tend not to be conditional, and there is less emphasis on graduation.
504. With such divergent, even polarised, views, it is all the more important that rigorous evidence should be made available to national governments, who, after all, are accountable to their citizens for the social protection programmes that they provide, and who should therefore be the ones taking full responsibility for their design and implementation.
505. Key contextual factors are: (i) political economy, (ii) affordability, and (iii) capacity. It is these, more than the impacts and technical design issues described above, that shape social protection in a country. For the first, there is considerable evidence that Government ownership is a key determinant of scale, success and sustainability. Successful programmes in Latin America, South Africa and India were all originated by Government and operated at national scale. NGOs, with funding and technical assistance from donors, have often implemented pilot programmes designed to demonstrate the beneficial impacts of social transfers. These have not generally found success with governments.
506. Affordability is closely linked to political economy. The financing of social protection programmes is much more a question of the political will to pay for them, than to any practical fiscal constraint. Even capacity in the end comes down to the Government's will to implement as much as it does to the availability of resources per se. Ethiopia's Productive Safety Nets Programme went from scratch to over 7 million beneficiaries in just a couple of years, using predominantly Government channels. All this suggests that it may be more important to undertake research into these contextual factors, to get a better understanding of the ramifications of the political economy, the determinants of fiscal space and the constraints of national capacity than to focus predominantly on impacts and technical concerns.

Assessment

507. The sector specialist found a ‘slight lack of coherence’ in the social protection portfolio. This was not surprising for the four grants through the open windows, although he noted that all four focussed on middle income countries, two of them Indonesia. However, even within SPTW the opportunity was perhaps missed to provide greater portfolio coherence. The scoping study was a good first step towards a conceptual framework and it developed some sensible proposals for research areas to be covered in the call:

OVERALL IMPACT - links between social protection and poverty reduction, asset accumulation, local economic development and community wellbeing with priority to: identifying context and implementation issues; and cost effectiveness/cost benefit analysis.

POLITICAL ECONOMY - how social protection programmes are financed, how to ensure sustainability, what mechanisms ensure the strengthening of overall social protection systems.

DESIGN – the most effective complementary programs/services, the most effective exit/graduation strategies, and the most effective targeting mechanisms. Cross-country and cross-regional studies to be encouraged.

The reviewer found that the ten SPTW studies fell short in each area:

- A number looked at impact, but very few considered long-term effects. Two studies on Ethiopia’s Productive Safety Nets Programme, using datasets going back to 2006 were the only ones 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 focussed 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, although one looked independently at four similar programmes in four different countries. More than half were studies of relatively small-scale, even pilot programmes, with limited likelihood of gaining the political support for scaling up. (Peru’s pension pilot being the most likely possible exception.)

508. Design issues were not emphasised. One or two studies compared alternative delivery channels, and one included an innovative delivery mechanism. Some considered the issue of exit/graduation and the interface with complementary programmes, but none compared targeting options or looked at the significance (or not) of human capital conditionality, two key areas of contention in the current social protection debate. There were none of the recommended “cross-country or cross-regional studies”. The one application looking at more than a single country comprised separate sub-studies of little potential for cross-country learning. Although it should be noted that studies of very similar programmes in Zimbabwe and Malawi offer a useful pairing for comparison.

509. As a whole, the reviewer felt the SPTW portfolio seemed unbalanced. Only three studies look at national-scale social protection programmes, and these are all employment/public works schemes: PSNP in Ethiopia and MG-NREGA in India. A fourth, on apprenticeship training in Ghana, might be considered national, but the employment-orientation would not be considered social protection under many definitions. The remainder are evaluations of small-scale pilot projects, many of them essentially donor-driven. As discussed earlier, these lend themselves better to an experimental, RCT approach; but the contributions to knowledge and to policy influence are likely to be far lower.

510. The reviewer noted the concentration of research institutions: four grants to IFPRI and two to the University of North Carolina.⁴³ Two of the IFPRI grants study one Ethiopia project, and the two North Carolina projects study essentially the same type of project in the same way, and for the same purpose, in neighbouring countries. It could be argued that this is an economical way to build a solid body of evidence on critical issues. The reviewer, however, felt it will limit the potential richness of the research under the thematic window.
511. The reviewer concludes that 3ie's social protection portfolio is useful, but not dramatically so. More than half the studies under the SPTW are RCTs evaluating the short-term impacts of small-scale pilot programmes. This ignores the realities of social protection: that proven impact is no guarantee of political support. Nor is it evidence of cost-effectiveness at a larger scale. In the worst case, the end result will be valid proof of positive impact, and a pilot that ends leaving its temporary beneficiaries no better protected than they were before.
512. The thematic window, in particular, could have made better use of the valuable scoping study that preceded it to ensure a better balance, in terms of geographical coverage and research methods, but above all in terms of the type of programme covered and the focus of the research.
513. The reviewer notes that the scoring criteria, and the standard application and grant review formats used for SPTW made no reference to specific research questions or to the areas identified in the Scoping Report. 3ie's Request for Proposals set out a broad definition of topics that would be considered for grants under the window:

...."risk reduction/coping mechanisms (in particular cash transfer interventions and public works), but may also consider complementary interventions in the area of risk prevention (unemployment benefits, pensions, and insurance) and opportunity promotion initiatives, such as job creation and training. This will enable the call to capture the linkages between the "protection" and "promotion" roles of social protection which fills an important evidence gap in this emerging field."

A tighter specification and a requirement that grantees justify the relevance of their proposal to the areas identified in the Scoping Study might have resulted in the selection of a portfolio with greater overall coherence.

6.4.3 Conclusion

514. Unless policymakers and implementers are ready to identify what issues they want to see evaluated, Open Windows will inevitably be driven by what interests researchers. However great the effort put into selecting grants for policy relevance, and reinforcing them with Policy Influence Plans, it is unlikely that further Open Windows will generate a solid body of evidence around a particular issue, and it will be a matter of chance if that issue is of high priority to policymakers.
515. It is, therefore, recommended that 3ie give greater emphasis to Policy Windows and Thematic Windows. Work is needed to develop the most effective way to operate both of these more strategic windows. PWs are expensive to run, but they are likely to generate a better understanding of the policy process than PIPs on OW grants, especially if they can be linked to efforts to increase 3ie's southern membership.
516. For thematic windows, a way needs to be found to ensure that proposals which respond to the strategic priorities defined in scoping studies, rather than just falling in the broad area of interest. If that can be done, they offer three important advantages:

⁴³ The impression of concentration may be reinforced by the fact that an IFPRI staff member is named as External Peer Adviser on the two North Carolina grants and another IFPRI staff member is Lead PI on one grant and EPA on another.

- as part of the scoping exercise, 3ie will identify in advance likely theories of change and key context factors, allowing stronger assessment of these aspects of the grants proposed
- provided that themes are selected carefully and well defined, policy relevance should be a given
- there are likely to be fewer, more capable grant applications, in particular with respect to their external validity.

6.5 Conclusions

517. This section reviews 3ie's relevance and effectiveness at four different levels, working from the bottom of its current Theory of Change, up to the top:

1. Recognition, Funding and Advocacy for Impact Evaluation
2. The Generation and Synthesis of Evidence
3. Building Capacity
4. Influencing Policy and the Use of Evidence

518. It is far too early to attempt an assessment of the impact, even the potential impact, of 3ie's work on the highest level of the ToC: Enhanced Development Effectiveness/ Better Development Outcomes. Suffice to say that any effort to support these outcomes must be highly relevant.

6.5.1 Recognition, Funding and Advocacy for Evaluation

519. The bottom level of 3ie's Theory of Change has two elements: '3ie is known and recognized' and 'funding is secured'. The two are linked by 'Advocacy for Impact Evaluation'. Both the Stakeholder Survey and the Evaluation Team's consultations show 3ie to have been highly effective in winning recognition for its technical capability and the quality of its services. While many others have contributed, notably the Center for Global Development, the World Bank, and the US researcher community more generally, 3ie has also played an important part in making the case for funding to fill the Evaluation Gap. There are now a good number of large, well-funded initiatives supporting Impact Evaluation: DIME and J-PAL/IPA are only the most well-known, each with over 300 IEs. Even compared with these large organisations, 3ie itself makes an important contribution, with 78 IE grants awarded, and more in the pipeline.

520. The relevance of 3ie raises two questions. One has been hanging since the Theory of Change was first drafted: recognition for what or as what? There are a number of possibilities: as a funder/manager of IEs; as an advocate for IE; as a builder of capacity for IE; or, as a standard setter in IE. While 3ie has won recognition in all these areas, the Evaluation Team believes it is the last which defines 3ie's most important role. Other organisations are as well funded and in most respects as well equipped to manage impact evaluations. Other organisations have a clearer mandate for capacity building. No other organisation has given the same importance to setting out what constitutes rigorous, policy relevant impact evaluation.

521. As far as advocacy is concerned, it seems clear that the case for IE has been made. 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. 3ie publications and the two major conferences have made important contributions to this. 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. In essence this

is 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.

522. The Evaluation Team believe 3ie is the only organisation playing this role. It certainly 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 which gives priority to statistical rigour and experimental technique above all else.
523. The success which 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. That is what 3ie's membership structure is intended to achieve. There is a tension between this and the fact that the largest share of funds available for evaluation are tied to particular sectors, and that decision makers give priority to sectoral concerns, often relatively short-term concerns of interest to a small group of stakeholder institutions. As 3ie is beginning to find, it is relatively easy to win funding for dedicated, sector-specific Thematic Windows and Systematic Reviews on particular issues. Convincing donors to provide core funding for real public goods, setting standards and managing high quality studies of enduring questions, is much more difficult. Nevertheless, the Evaluation Team believes it is vital that 3ie makes the effort to maintain its higher, public-good function. Otherwise it risks becoming a service provider which, however effective, must tailor its strategy to the availability of dedicated, mostly sectoral funding.

6.5.2 Generation and Synthesis of Evidence

524. The two elements of the mid-section of the ToC are currently: 'rigorous IEs are undertaken', and 'Existing evidence is analysed and synthesised'. Each is underlain by a major assumption. The first is that if funding is available, rigorous IEs will be undertaken. The second is that high quality evidence is available for analysis and synthesis. If there is one central conclusion from the review of the potential impact of 3ie grants, it is that neither of those assumptions will hold until a number of other conditions are met.
525. In trying to generate rigorous, policy-relevant evidence, 3ie has faced the two classic dilemmas of grant management. The first is that if selection standards are set too high, there may be no grants worth funding. The second is that once a grant has been awarded, the effort required to extract an acceptable result can become so great that it might have been better to do the work in-house. To the extent that 3ie is undertaking SRs in-house, this has actually happened. 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 may be disappointing. There is no reason to believe that other organisations have done better. It seems clear that 3ie processes are more rigorous than most. 3ie PIs include many of the most respected IE specialists, a number of them are senior staff in those other organisations. To take one well-known example, 3ie's experience with cook stove IEs closely mirrors that reported by J-PAL.⁴⁴ To the Evaluation Team, they are both examples of the incomplete Theory of Change and inadequate reconnaissance effort identified above as common weaknesses.
526. Various reasons can be suggested for these problems. It was perhaps inevitable that a rapid increase in funding for IE would stretch the available capacity, even in developed countries. The result is that some of the better known practitioners seem to be overstretched. Apart from medical trials, the most rigorous experimental techniques have been developed by behavioural economists working in developed countries. This may have resulted in too narrow a focus on

⁴⁴ See 3ie's OW1.69 and www.hindustantimes.com/News-Feed/AbhijitBanerjee/Not-fired-with-logic/Article1-895712.aspx

this set of theoretical tools. There are also many lessons to be learnt about how to apply rigorous experimental techniques in the diverse, rapidly changing and complicated circumstances faced by poor people in developing countries. The implementation challenge has clearly been underestimated. 3ie's portfolio includes many where the Lead PI only spends one month a year in-country, with supervision largely delegated to less experienced national colleagues. Above all, few studies have sought to base their theory of change on a proper understanding of existing theoretical and practical literature. There are few sectors where this is not both substantial and closely argued.

527. Quality issues of this kind 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 structure behind it. Nevertheless, the Evaluation Team believes there is considerable scope to raise the standard of impact evaluation for development. Many of the weaknesses identified are matters of simple good practice. Robust quality assurance and advocacy for good technique has the potential to raise standards significantly. 3ie's own reviewers identified quite a few of the same issues as the Triple Line reviewers and the Evaluation Team feel they could have followed through more strongly with the grantees. If 3ie were to have the courage of its convictions, on theory of change, on mixed methods and on cost effectiveness, it would make a great difference. However, this can be difficult, if it means confronting senior academics from well-known institutions. For this reason, the Board has an important role to play, giving 3ie clear guidance on how high to set the bar, and when necessary the backing they will need to enforce it. A procedure to bring cases of poor performance in front of the Board, might be appropriate.
528. 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.

6.5.3 Building Capacity

529. It is understood that 3ie was never expected to have a major capacity building role. The ToC element 'capacity to undertake IEs and SRs' is principally expected to come through developing country researcher involvement in the implementation of IEs working with an international Lead PI, or with the support of international consultants. Over time, 3ie has strengthened its requirements, asking grant applicants to state the involvement of developing country researchers and how their capacity will be built. What is proposed ranges from pure on-the-job training through to attendance at a formal course overseas. 3ie does not suggest that a budget be proposed for capacity building, and in the proposal review capacity building is only scored as a sub-set of developing country involvement.
530. There are number of grants where there is a strong relationship between national and international research institutions working on a 3ie grant, especially in China and Mexico. However, these relationships were established before the 3ie grant was awarded.
531. It is not possible to assess directly how effective 3ie has been in building capacity, but the overall impact is unlikely to be large. This need not be a criticism. Other organisations are better equipped for this task, and have both the mandate and the funding. The CLEAR initiative is the most prominent example (Regional Centers for Learning on Evaluation and Results).

6.5.4 Influencing Policy and Use of Evidence

532. Improving the use of rigorous evidence in policy and programming is widely agreed to be an important contributor to more effective policies, delivering better outcomes for poor people in developing countries. The problem is that this is not an evidence-based statement: the theory of change between ‘better evidence’ and ‘better outcomes’ is also agreed to be complex and very likely context-specific. There are on-going arguments in the literature and between practitioners about what ‘evidence-based policymaking’ means in practice. Some argue that getting the politics right is in fact the most important precursor to making effective and sustainable policies⁴⁵. Rigorous evidence contributes to, but does not drive, better policymaking. Process maps of the policy process drawn for 3ie’s Rome clinic in April 2012 show that there is a complex network of national and international organisations, working groups, technical committees, advisory groups and planning and budgeting processes involved in policymaking. This is not limited to developing countries: it is a facet of modern policymaking that reflects the complex and often overlapping nature of the challenges facing policymakers and the different types of evidence being used⁴⁶.
533. It may be relatively clear what evidence is relevant to policy: the evidence covers the issue in question and may confirm hypotheses, challenge received wisdom, explain complex relationships or simply enrich our understanding. But this does not mean that it will necessarily have an impact on policy. Policy influence is conditional on the interactions between different groups and the quality of the policy process, not on the position of a person or organisation in the hierarchy or network. Good evidence put through poor policymaking processes will result in poor policies. And policy proceeds on the basis of the weight of evidence, particularly where the policy is a new one. A single large study may change how a current policy is implemented or funded, or may alter the extent of its coverage or distribution between different groups. A study on an issue that is not covered by an existing policy, such as some of the proof-of-concept studies currently funded by 3ie, may come up in debate in these groups but is unlikely to affect change unless it is bolstered by other evidence. Whether this counts as influencing policy is debatable.
534. The first challenge for 3ie is to work out how far its remit extends along the theory of change between promoting the use of better evidence and changes to policy and programming intended to deliver better outcomes for the poor. The second challenge is to determine whether such a theory of change is a universal one and therefore whether 3ie has (by virtue of encouraging its grantees) a general remit to influence policy. Neither of the PACO strategy documents explores these issues. While 3ie’s founding document talks about ensuring the policy relevance of the evidence it promotes, the Learning and Advocacy Strategy discussed policy relevance and policy influence in the same section, inferring one from the other. This has resulted in a degree of mission creep, as discussed in section 6.1. In OW1, grantees were encouraged to develop a communications strategy for their research. For some, this involved ensuring that key policymakers were informed and for one project (OW1.66) the feedback from their dissemination activities was that the project’s evidence would be fed into the Chinese policymaking process. This project was used to pilot PIPs, giving an erroneous picture of how policy influence could be assessed. Chinese policymaking – as far as can be determined – follows a very different model to other countries.
535. Survey results show that some PIs are comfortable with attempting to use their evidence to influence specific policies. Others are not. The question for 3ie is whether encouraging grantees to advocate for particular policy positions based on evidence from a single study is a step too far

⁴⁵ See Hallsworth, M 2011 *Policymaking in the real world: evidence and analysis*. London: Institute for Government.

⁴⁶ See Jones et al 2012 *Knowledge, policy and power in international development: a practical guide*. Bristol: The Policy Press.

from its original remit of promoting and financing impact evaluations that ‘address questions that are relevant and important to public policy decisions’. The evaluation team believes that it is. The theory of change between promoting rigorous evidence and policy change is not a universal one and in any case is too weak for policy influence to be a central objective for all grantees.

536. It is important for 3ie to remain a trusted source of independent evidence on what works and why, identifying the most pressing questions relevant to policy and then adhering to the highest possible quality standards in attempting to answer them. It also means setting the bar to encourage others to do the same. While there is always a need to communicate research results widely and well, the evaluation team’s conclusion is that 3ie needs to focus on ensuring good research communications, and to rein in the current enthusiasm to demonstrate policy influence.

6.5.5 Development Outcomes

537. It is too early to identify any attributable impact on development outcomes from 3ie’s work. Its work is directed at some the most pressing issues in development, especially in the social sectors. There is every reason to believe that better evaluation evidence on those issues will strengthen the drive towards better policy and more effective implementation. As set out in the discussion at the beginning of this chapter, there are two critical assumptions: that the political and fiscal constraints are not too great to allow policy makers to use the evidence they are given; and that the evidence is both rigorous, in the broader sense proposed here, and credible. It is the second of these assumptions which provide 3ie with its reason for being.

7 Conclusions and Strategic Review

538. The Evaluation Team have had extensive access to 3ie's records and open consultation with a broad range of stakeholders. They have been able to make a detailed review of a reasonable sample of 3ie outputs. They have seen sufficient evidence to present these conclusions with confidence. It is important, nevertheless, to stress that it is impossible for any evaluation to capture every aspect.
539. The conclusions are presented under two headings: Process and Relevance and Effectiveness. After noting achievements under each heading, a distinction is drawn between relatively minor issues and those which have more importance for 3ie's operation and strategic direction.
540. It is important to state the 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.
 - 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 Process

541. In the three years since it was incorporated, 3ie has made very substantial progress in a lot of areas: 68 IEs and 15 SRs commissioned, through rigorous and transparent selection processes and managed to a high standard. A strong, enthusiastic team has been established, and they have worked together to build and implement efficient processes in all areas of work. The working relationship between the Board of Commissioners and 3ie management now operates effectively.
542. One minor issue merits attention: procurement and management of major non-grant contracts, IT contracts in particular. A rather more important issue concerns performance management and HR, always critical for a start-up organisation. 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 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.
543. There are three more major issues. In order of importance they are: (i) cost; (ii) planning, budgeting and reporting; and, (iii) membership.
544. The Evaluation Team are wary of a conclusion that 3ie's administrative costs are high relative to other grant funding agencies. It is vital to ensure that like is being compared with like when benchmarking these costs. That said, the GDN hosting arrangement seems poor value for money and it can be argued that 3ie's institutional costs are high. The distributed structure across Delhi,

London and Washington does impose additional cost. It is difficult to see an effective alternative. There may be a case for moving some functions from London and Washington to Delhi, but the Evaluation Team have not identified any obvious candidates.

545. 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.
546. 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 of relevance and effectiveness, 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 to this issue.

7.2 Relevance and Effectiveness

547. The extent of 3ie's support from a diverse group of stakeholders is clear recognition of the relevance and effectiveness of what it has achieved to date. The body of clear, authoritative material it has generated on the principles and practice of impact evaluation is another vital achievement, together with progress in adapting the principles of systematic reviewing to the more diverse context and varied evidence base of international development. 3ie has played an important standard-setting role.

548. It is, however, in the area of Relevance and Effectiveness that 3ie faces its most major challenges. There are four, summed up under the headings: Policy, Quality, Relevance, and, once again, Membership.

(i) Policy

549. 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 placed too much emphasis on the last, in particular on the mechanisms of influence.

(ii) Quality

550. This leads to the second point, quality. 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 strong set of 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

551. 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. 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 degree of focus set out in the scoping study, but experience should improve the process. The tendency towards a menu approach, where funders select questions of interest within a broad theme 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

552. 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. For southern members, the Evaluation Team suggest a broader approach. The aim would be to bring national policymakers and evaluation practitioners together to brief them on theory-based impact evaluation and assist them to review how it may fit into their overall M&E effort. This would be an extension of the demand generation workshops, but the emphasis would be on encouraging membership and national demand for evaluation, not necessarily leading to immediate IE grants through the Policy Window. The Evaluation Team recognises that this will not be easy, and that it will require a targeted, and funded effort to achieve.

553. This leads to the second recommendation, that 3ie should make an immediate start on drafting a Strategy 2014 – 2017, and with the Board's support, plan a campaign to enlist donor members willing to provide the core funding needed to implement it. The strategy should emphasise the public good nature of the work to be supported: in setting standards, in advocating for theory-based evaluation and in enabling southern members to become active participants in all aspects of impact evaluation.

7.3 Strategic Review

554. This section briefly sums up some suggestions 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 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.

555. 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. It will also reduce the scope for the 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 programme along the lines described, and campaigns to win core funding for the public goods it will provide under the programme.

List of People Met

(in alphabetical order)

Oliver Babson, Deputy Director for Strategy, Planning & Management, Bill and Melinda Gates Foundation

Tania Barham, Assistant Professor Economics, University of Colorado at Boulder

Lori Beaman, PI, OW3.1236, North Western University

Ami Bhavsar, Research Assistant, 3ie

Annette Brown, Deputy Director - Advancement and Impact Evaluation Services, 3ie

Lorenzo Casaburi, Doctoral Research Fellow in Sustainability Science, PI OW3.1060, Kennedy School of Government, Harvard University

Christelle Chapoy, Advocacy and Policy Officer, 3ie

Stacey Davies, Journals Managing Editor - Development Studies, Politics & IR, Taylor and Francis

Thomas de Hoop, Evaluation Officer, 3ie

Geoff Deakin, former Director, Sydney Medical School Foundation

Olivier Durand, Programme Coordinator, World Bank Malawi

Paul Fatch, Researcher, Malawi Ministry of Agriculture and Food Security

Dugan Fraser, Associate Researcher, Economic Policy Research Institute

Marie Gaarder, Director, Evaluation Department, NORAD

Sulley Gariba, President, African Evaluation Association

Paul Gertler, Professor of Economics, Haas School of Business, University of California, Berkeley.

Alison Girdwood, Education and Results Adviser, DFID

Adam Handler, Chief Policy, Advocacy & Communications Officer, 3ie

Gonzalo Hernández Licona, Executive Secretary, Mexico's National Evaluation Council (CONEVAL)

Katie Hughes, Senior Grants Administrator, IPA

Bindu Joy, Administrative Assistant, 3ie

Jatin Juneja, Finance Officer, 3ie

Sullay Kamara, Policy counterpart, OW3.1180, Centre for Economic and Social Policy Analysis (CESPA)

Jessica Kiessel, Director of Country Programs, IPA

Jeff Knezovich, Policy Influence and Research Uptake Manager, Institute of Development Studies

Manveen Kohli, Program Manager—COHESIVE, Duke Global Health Institute

Shari Krishnaratne, Evaluation Officer, 3ie

Randall Kuhn, PI, OW3.1060, Duke University

Arianna Legovini, Advisor and Head, Development Impact Evaluation Initiative, World Bank

Ruth Levine, Director, Global Development and Population, Hewlett Foundation

Emmanuel Lyimo, Field Monitoring and Evaluation Coordinator, Tanzania Forest Conservation Group

Minna Madhok, Programme Associate, 3ie

Thabit Masoud, Policy counterpart, OW3.1109, Care Tanzania

Debbie Menezes, Programme Manager, DFID, India

Jane Menken, Director, Institute of Behavioral Science (IBS), PI 3.1060, University of Colorado at Boulder

Radhika Menon, Communication Officer, 3ie

Sunita Menon, Director Education & Training, Breakthrough

Charles Meshack, Executive Director, Tanzania Forest Conservation Group

Ashima Mohan, Programme Associate, 3ie

Paromita Mukhopadhyay, Online Marketing and Communications Officer, 3ie

Nyokabi Musila, Research and Knowledge Translation Scientist, African Institute for Development Policy

Ghulam Mustafa, Director PEC, Earthquake Reconstruction & Rehabilitation Authority, PM's secretariat, Pakistan

Andrew Nyandigisi, Focal Point Case Management of Malaria, Ministry of Public Health, Kenya

Arnaldo Pellini, Research Associate, ODI

Lauren Persha, PI, OW3.1109, University of Michigan

Jyotsna (Jo) Puri, Deputy Executive Director and Head of Evaluation, 3ie

Laura Rawlings, Lead Social Protection Specialist, World Bank

Nafis Sadiq, Special Advisor to the UN Secretary-General and Special Envoy for HIV/AIDS in Asia and the Pacific, UN

Anita Sircar, Programme Officer, 3ie

Birte Snilstveit, Evaluation Officer, 3ie

Hitesh Somani, Deputy Director - Finance and Administration, 3ie

Mukul Soni, Administrative Assistant - Communication, 3ie

Lyn Squire, Ex-President, GDN

Alan Stanley, Senior Thematic Convenor, Institute of Development Studies

Mathilda Stevens, Chief Director Monitoring and Evaluation, Department of Social Development, South Africa

Alvaro Velezmoro, Programming and Strategic Planning, CEPLAN, Peru

Arun Virk, Programme Officer, 3ie

Hugh Waddington, Senior Evaluation Officer, 3ie

Howard White, Executive Director, 3ie

Chris Whitty, Chief Scientific Advisor, DFID

Nick York, Head of Evaluation, DFID

Chowdhury Zafrullah, Trustee, Gonoshasthaya Kendra Trust

Abebe Zerihun, IPA country director, Sierra Leone, IPA