IMPACT ASSESSMENT

AN ONLINE SPEAKER'S CORNER DISCUSSION LED BY GARY WOLLER AND HOSTED BY microLINKS.ORG

September 26-28, 2006

CONTENTS

WELCOME
Post By Gary Woller

Hello Speaker's Corner Participants,

I would like to welcome everyone to the Speaker's Corner on impact assessment. The specific context of the Speaker's Corner is the impact assessment of so-called Private Sector Development Programs (PSD), which are programs aimed at increasing economic growth and wealth creation in poor communities via, for example, facilitation of MSE linkages to more lucrative and rapidly growing markets and through provision of an array of supporting services, including inputs provision and finance.

Before we launch into the discussion, I think it will be useful to place impact assessment within an appropriate context. A central assertion of impact assessment advocates is that the ability to attribute outcomes to program operations has both inherent and practical value that go beyond the ballpark guesses intrinsic to alterative evaluation methodologies.

Nobody who advocates impact assessment would argue that every program needs to be assessed nor that it is even necessary to assess a majority, or even significant percentage, of programs. Bur, taking a more strategic approach to impact assessment arguably addresses many of the practitioner criticisms, although perhaps not totally mitigating them.

I would like to address these and other issues in this Speaker's Corner. My primary interest is jointly investigating with participants how impact assessment can be made more relevant to their work in private sector development. This, of course, involves a discussion about purpose, strategy, methodology, benefits, costs, and the like.

I look forward to your comments, and I hope that I can address them satisfactorily.

DAY ONE: DEFINING IMPACT ASSESSMENT
Post By: Gary Woller

What can Impact Assessment do and not do? How does it differ from other forms of program evaluation? What is the role and potential role of impact assessment in program planning and management?

RE: DAY ONE: DEFINING IMPACT ASSESSMENT
Post By: Dan Norell

Gary,

There has been a lot less use of impact assessments in microfinance in the past years. The cost is just too high for the MFIs. Donors have been requiring them less and less.

The gain is the less cost to the MFIs. The loss is that we had less of an ability to dialog with other development practitioners from other sectors about the positive impact that microfinance and enterprise development has on clients. We argue that it does, but I am do not hear donors or other practitioners stating that microfinance and enterprise development has a positive impact on the lives of the poor.

RE: DAY ONE: DEFINING IMPACT ASSESSMENT
Post By: Lara Goldmark

There is a big difference between the two areas you mention, microfinance and enterprise development. In microfinance, whether or not they are recent or ongoing, there is a critical mass of credible research out there that shows that microfinance does have a positive impact on the lives of clients - albeit in ways slightly different from those that might have originally been envisioned, for example the impact of microfinance is more related to income stability and protection from shocks and there is much less evidence about the growth of client enterprises.

On the enterprise development side, credible impact assessments are few and far between - it doesn't help that there are so many different types of enterprise development projects, or these days with the integrated approach, so many interventions within one intervention - we are a long way away from having a critical mass of credible research.

RE: DAY ONE: DEFINING IMPACT ASSESSMENT
Post By: Mike Albu

Dear Gary and colleagues

Program staff from Practical Action convened recently to improve IA systems in our Markets and Livelihoods programme - which works for pro-poor market development in very diverse contexts in Africa, South Asia and Latin
America. We had some input from WISE, the UK consulting group responsible for the DFID-funded Enterprise Development IA Information Service (www.enterprise-impact.org.uk). The following hierarchy of terms proved useful for our discussions, so I offer it here as a way to kick off our conversation:

Project / Program Monitoring:

The regular systematic collection and analysis of data and information about inputs, activities and outputs - for management purposes: i.e. to monitor progress against a plan, identify emerging issues and problems, take corrective action.

Performance Evaluation:

A periodic study at key points in the project cycle, that examines more broadly the performance of a program / organization - e.g. looking at staff behavior, organizational structures and methodologies, the efficient achievement of outputs. Performance evaluations should assess the effectiveness of different factors which are being used to effect change; and create opportunities for implementing organizations to learn about themselves.

Impact Assessment:

IA is about outcomes: what happened? Impact assessments examine changes at the higher level objectives of a programme - ie. the goal and purpose level. Their effectiveness of IA studies depends in part on having a baseline or a
control group to compare results with - so that changes can be properly attributed. IA should be able to identify what factors contributed to the changes, and how. Impact assessment exercises are also key opportunities for both program staff and beneficiaries to reflect upon and inform evolving program design.

Put like this, it is obvious that effective IA depends on good PE and in turn on systematic PM.

However, it is also surely important that Impact Assessment leads, via the learning it catalyses, to responsive changes in the monitoring systems and performance measures that we employ.

Regards

RE: DAY ONE: DEFINING IMPACT ASSESSMENT
Post By: Zan Northrip

This is a great question, because the semantics really matter.

Project evaluation often boils down to an evaluation of whether or not a contractor has met the requirements outlined in its contract. There is a heavy focus on project execution, as opposed to project impact -- did the personnel show up on time, did they submit the promised deliverables, etc. Industry parlance confirms this; we usually say "So and so (the
contractor) got a positive evaluation on that project." While I think there is a place for project evaluation (someone has to hold us contractors accountable!) these kinds of evaluations don't tell us much about whether a project's design made any sense, or whether the project's interventions actually led to any attributable results. We need impact assessment to shed light on those latter points, and to guide us on how future projects should be shaped to achieve the highest return on investment.

I think if we can make the distinction between project evaluation and impact assessment repeatedly and consistently, we stand a good chance of increasing the willingness of donors to dedicate more resources to impact assessment.

RE: DAY ONE: DEFINING IMPACT ASSESSMENT
Post By: Gary Woller

Dan:
I agree that for the "typical" MFI, the cost of doing impact assessment (that is, high quality impact assessment) is probably too high, except in certain cases (e.g., large or innovative program with abundance of donor funding or significant international notoriety). My experience with MFIs is that they are more interested in simpler, evaluation methods that are at best suggestive of impact, but which can feed into the program's operating and decision making cycle. If this is the objective, impact assessment may not be the best alternative. (Note that by the term "impact assessment," I
mean a high-quality assessment with valid control groups--one that meets the minimally acceptable methodological standards suggested by the PSD Impact Assessment Initiative.)

That said, "strategic impact assessments," of selected programs, with high potential for learning or generalization, is entirely appropriate. Strategic assessments are probably best funded by donors with minimal involvement by the program itself (except where necessary). The purpose would be not be necessarily to assess the program itself, per se (though it can be a purposeand is unavoidable), but to assess the methodology, a general approach, an innovation, etc.

Though there have been literally dozens of impact assessments of microfinance programs (probably over 100-- I've documented the methodology of over 80 of them), few, however, have satisfied the minimally acceptable methodological standards, thus it would be of value to conduct further impact assessments of microfinance programs using high-quality methods. This would not only allow you to engage with other development sectors, but it would probably give you a leg up on them, since so few quality impact assessments are being done elsewhere in other sectors (at least as far as I know).

But since strategic impact assessments constitute a public good, the issue of funding, and interest among donors, becomes crucial. I cannot account for the apparent lack of interest by donors mentioned by Dan, perhaps someone else can comment on it.

It would be interesting to hear what others suggest in terms of getting donors more engaged and interested in doing impact assessments of microfinance and other PSD programs.

Gary

RE: DAY ONE: DEFINING IMPACT ASSESSMENT
Post By: Dean Karlan

Dan:
I would suggest that much of the drop off has been because the impact assessments do not provide enough tangible prescriptions to MFIs. This can be resolved by being more creative in the design of impact assessments. Too often an impact assessment measures only credit vs. no credit. Impact assessments designed creatively can ask the relative impact of two approaches (relative to each other, or relative to a control group). This helps provide tangible prescriptions, informs stakeholders on the impact of their dollars spent, and helps them direct future dollars to approaches that maximize future impacts.

RE: DAY ONE: DEFINING IMPACT ASSESSMENT
Post By: Gary Woller

Lara makes a good point. As I pointed out to Dan, there are probably in excess of 100 microfinance "impact assessments," though the number of "high-quality" assessments is a small fraction of the total. (This is similar to what the PSD Impact Assessment Initiative found in its review of impact assessments of PSD programs. See "Review of Evaluations of Selected Enterprise Development Projects" by Zandniapour, Sebstad, and Snodgrass, which can be found at the PSD Impact Assessment Initiative website.)

Generally, the evidence suggests there is an impact of microfinance programs, although the precise impact varies considerably and is very hard to generalize, while many of the results lack credibility. Perhaps some more high-quality assessments would help to clarify this particular information gap.

Based on the work of the PSD Initiative, we conclude that we do not know much about how effective PSD programs are, such that there does appear to be significant value to investing resources in a strategic manner to determine what this impact is for different types of intervention strategies.

Gary

RE: DAY ONE: DEFINING IMPACT ASSESSMENT?
Post By: Elizabeth Dunn

Gary, you pose some important questions for the first day of discussion.

I think that the definition of terms provided by Mike Albu of Practical Action is useful and helps to answer the second question: How des it (impact assessment) differ from other forms of program evaluation? We should keep these distinctions in mind during the discussion.

Similarly, Lara Goldmark of DAI makes an important point about the differences between microfinance and enterprise development. And, while we have accumulated some credible evidence about the impacts of certain microfinance products, there is much less evidence about the impacts of enterprise development programs. Enterprise development programs, particularly to the extent that they are embedded in more integrated programs for private sector development (PSD), present some special challenges to impact assessment (which we may end up discussing later).

But back to today's question: What can Impact Assessment do and not do? What is the role and potential role of impact assessment in program planning and management? I think that the first posting by Dan Norell of World
Vision points to one of the most important potential uses of impact assessment: impact assessment can inform resource allocation. Dan mentions dialogue within the community of donors and practitioners, all of whom are looking for successful poverty alleviation programs.

Without impact assessment, how do we know what is successful and what isn't? How should the scarce resources (money and dedication) be allocated? I say that impact assessment has the "potential" to improve resource allocation,
but that depends on several conditions 1) the results are valid and credible; 2) there is some degree of comparability across impact assessment results of different types of programs; and 3) perhaps hardest of all, we actually take the time to compare and interpret the results across various types of programs. Those are all big "ifs".

RE: DAY ONE: DEFINING IMPACT ASSESSMENT
Post By: Gary Woller

Dean:

Thanks for the response. Your explanation rings true to me given my experience.

Knowing whether a program has impact and what the impact is certainly valuable. More valuable yet is this information combined with practical prescriptions/recommendations for policy planning and program operation ismore valuable yet.

Have you not been involved in conducting the types of impact assessments to which you refer? Could you give Dan and others some specific examples along with citations (and web links if they exist)?

Demonstrating the practical utility of impact assessments is, I think, critical in convincing skeptical audiences of why they should invest in them.

Gary

RE: DAY ONE: DEFINING IMPACT ASSESSMENT
Post By: Amy Davis Kruize

Dear Colleagues,

I do think there are some significant studies on impact of microfinance, which we refer to quite readily when looking for funding, but the ones of high quality are few and far between. Many MFIs don't have the in-house expertise to easily facilitate quality impact assessment, as well as the time and money to do so, thus have turned to more practical and periodic use of client assessment tools-- i.e., client satisfaction, etc. to tell them how they are doing.

From my stand point, good impact assessment is quite costly for an emergent MFI, which is struggling to achieve self-sufficiency, avoid delinquency and compete with the ever growing number of other financial service providers.

So--an interesting question to ponder might be a) who really needs and uses impact assessment data? and b) who's going to pay for it?

Amy

RE: DAY ONE: DEFINING IMPACT ASSESSMENT
Post By: Gary Woller

It is entirely feasible that a program could be designed poorly, with an inappropriate causal model, and with minimal to no impact but still receive a high evaluation score. Doing good impact assessments is as much about asking the right questions as anything else.

Whether a contractor meets its requirements, whether the program operation kept within budget, whether service delivery was good, whether the program met its stated and measurable input and output objectives, these are all valid questions, but they do not answer the fundamental question: "What is the impact of the program?" Nor can they answer questions related to critical causal relationships and their implications for program operation and effectiveness.

A problem that continues to arise is that of terminology. For example, in the review of evaluations of PSD programs I referred to earlier, researchers and authors consistently use the term "impact assessment," even though their study makes no attempt, or a highly deficient attempt, to establish attribution. Impact assessment, in this context, really refers to "program evaluation" of various types. While we cannot necessarily, at least in the short term, change the way thousands of people define "impact assessment," we can be consistent in our own use of the term and be clear about what it means.

We (that is, the PSD Impact Assessment Initiative) propose that the term "impact assessment" be used to describe program assessments that satisfy a set of minimally acceptable methodological standards, which are:

1. It will include observations on a group of program participants (treatment group) and a matched group of non-participants (control group).

2a. If quasi-experimental methods are used, it will assess the status of both treatment and control group members at a time before impacts can have occurred (baseline) and at a time after impacts can reasonably be assumed to have occurred (follow-up).

2b. If experimental (random assignment) methods are used, it will assess the status of both treatment and control group members at a time after impacts can reasonably be assumed to have occurred (follow-up).

3. It will be based on a causal (logical) model with clearly stated hypotheses linking program activities to expected impacts.

4. It will be rigorous in that all methodologies are well documented and their weaknesses identified.

5. It will use data collection methods that follow accepted good practice.

6. It will use analytical methods that are appropriate in that they match the type of data collected.

RE: DAY ONE: DEFINING IMPACT ASSESSMENT
Post By: Keith Kranker

I agree completely with your points. However, I believe the practitioners have a problem: they are "in the business" of helping people. However, a "valid control group," would in reality be a micro-entrepreneur who would otherwise qualify for a loan from the MF program.

To create a true control, they would need to (1) locate & survey potential borrows (often including all of the work of organizing lending groups, etc..), (2) evaluate their worthiness for a loan, (3) determine that they would really benefit from a loan, and then (4) refuse the loan anyway (only for the purpose of creating a control).

Thus, most MF programs have a worse problem than Ulysses: they would have already (figuratively) set anchor and rowed to shore BEFORE "tying themselves to the mast."

Keith Kranker

RE: DAY ONE: DEFINING IMPACT ASSESSMENT
Post By: Dean Karlan

I fear that folks aren't interested in a discussion too thick on methodological issues, so my apologies in advance.... but I do want to say that the four things noted below are not "needs" by any means. There are definitely ways, ways that we have employed (e.g., http://aida.econ.yale.edu/karlan/downloads/Derationing.Karlan-Zinman.pdf), that do not suffer from the problems noted below and yet still formed a valid control group through randomization. Another approach (different than the paper cited above) is through randomized program placement. We have done the initial setup of this in both India and the Philippines, but no papers yet (the rollout is underway). It simply requires an organization that is expanding but can't go everywhere all at once. Finding such organizations is not difficult since indeed rare (if ever) is the unconstrained organization.

To Gary's question, other examples of studies which compare two approaches (rather than measure the impact of the intervention versus nothing), include:

1) business training: http://aida.econ.yale.edu/karlan/papers/TeachingEntrepeneurship.pdf
2) specialized savings product: http://aida.econ.yale.edu/karlan/papers/FemaleEmpowerment.pdf and

http://aida.econ.yale.edu/karlan/papers/SEED.pdf

3) group versus individual liability (note this initial paper does not compare impact on the clients, but the design is such that one could... a later paper will do that... here we measure impact on repayment and client retention, and a little on social networks):

http://aida.econ.yale.edu/karlan/papers/GroupversusIndividualLiability.pdf

-- dean

RE: DAY ONE: DEFINING IMPACT ASSESSMENT
Post By: Nhu-An Tran

Thank you Gary for posing some great questions to launch this discussion on impact assessment.

It's true that donor has shown less enthusiasm for impact assessment in recent years. As donor funding becomes more limited, impact assessment has become less of a priority and we tend to rely on more project evaluation to measure performance and "success" of an intervention. While it's true that on the microfinance side, there is a more extensive body of literature on impact, there remains some lingering doubts about the validity of the data and the usefulness of the data to the MFIs. One of the useful outcome of the AIMS is that it helped the microfinance sector better learn about the characteristics of poor households, their vulnerability to shocks, and their needs for financial services. This has made MFIs more responsive to the demand side of the equation than before.

But as Dean Karlan mentioned, how do we make impact assessment more prescriptive and useful to MFI managers, donors, and program implementers? How do we know that the intervention that we chose, or the product/service that we are providing, is making a positive difference in the household/microenterprise?

Those are the issues that I hope we will explore in the next few days.

Many thanks,

Nhu-An

RE: DAY ONE: DEFINING IMPACT ASSESSMENT
Post By: Tony Pryor

One key concern affecting the utility of IAs relates to the question of purpose; is the intent of the assessment to allow donors and their legislators a way to judge program impact, for donors and implementers on the ground to assess impact of projects, or of financial institutions to assess the impact of their activities? All are

reasonable objectives, but often the purposes and definitions of success are interwoven and confused.

And one key problem is the question of timing. Just when should an IA be carried out? All too often assessments are carried out in conjunction with the cycle of project funding. This would be fine IF true impact were achieved coincidentally when project funding ended. And this indeed is the case in many instances. But in others, projects set the stage for broader systemic impacts. What can be assessed at the end of the project is the POTENTIAL for broader impact. Investments may indeed have focused less on impact but on getting systems, policies and procedures in place, and experimenting with approaches. I don't know micro-fi well enough to know if this problem affects impact assessment, but it's rare that a donor looks for impact far after funding is complete.

RE: DAY ONE: DEFINING IMPACT ASSESSMENT
Post By: Gary Woller

It appears that Tony has made the same distinction between generalized and targeted learning as I did in my last reply.

As for timing, a general guideline is that the assessment should be carried out as soon as possible in the program cycle (or in conjunction with program expansion) so as to capture a baseline of where program clients/participants and control group members were before the intervention or before enough time had elapsed after they stated to participate in the program.

It is also important to match the timing of the follow-up to a reasonably long enough period of time during which impacts might have manifest themselves. The PSD Impact Assessment Initiative team uses a two-year rule of thumb, recognizing that in many cases that two years is perhaps insufficient time to see the full benefits, but also recognizing that two years may be pushing it in terms of satisfying stakeholder demands for timely information.

It may be possible to continue to assessment after the program itself ends, depending on how the program is structured. If the program, for example, works through local implementing partners, the program may end, but the partners will continue to implement the market interventions after the program ends. This is the case in an impact assessment I am involved with in India. The assessment is scheduled for 2 years, but the PSD program itself may end prior to that time. The program works through local implementing partners, who have a long-term commitment to the market and the market interventions. So, even if the program ends, we can continue to work through the local implementing partners who remain active.

Let me add here also that determining issues of timing is one of the primary purposes of the "Evaluability Assessment," which is a preliminary assessment carried out by a research team to determine whether the program is an appropriate subject of an impact assessment. (The EA also looks at the causal models for the

program to determine whether they are sound, whether the program is designed appropriately according to the causal model, whether there is support for the IA among the program, its partners, and other stakeholders, in addition to other issues. The PSD Initiative has conducted a number of Evaluability Assessments. They are in the process of being posted to the PSD Initiative website.)

RE: DAY ONE: DEFINING IMPACT ASSESSMENT
Post By: Dean Karlan

Two thoughts in response to the above message:

1) An impact assessment is traditionally defined in program evaluation academic literature as asking "How did the lives of the participants change relative to how they would have changed had the program/intervention/etc not existed at all?" Measuring how lives changed is easy (or easier). The hard part of any and all impact assessments is credibly measuring how lives would have changed had the program not existed. The ability to measure this counterfactual convincingly is what separates good from bad methodologies.

2) "Their effectiveness of IA studies depends in part on having a baseline or a control group to compare results with..." I agree completely that the effectiveness depends on having a control group. However, a baseline is not necessary. I have found this to be a common misperception amongst practitioners, and hence wanted to raise it here. Randomization, and a sufficiently large sample, obliterate the need for a baseline (the randomization ensures that treatment and control are statistically similar on both observable and unobservable characteristics... and this can be verified on the observables just to make sure). There are of course nice things one can do with baselines, but it is not necessary for answering the basic impact assessment question defined above in #1, as long as you have a valid control group.

-- Dean

RE: DAY ONE: DEFINING IMPACT ASSESSMENT
Post By: Gary Woller

Amy:

Your experience has been mine as well, with regards to "in-house" impact assessments. I refer you (and others) again to Dean's comments about how to make impact assessments more relevant for practitioner programs. I assume (hope) he'll elaborate on his points a bit more.

As for who needs IA and who's going to pay for it, I think to answer this, we need to be clearer about what it's purpose is or purposes are. Perhaps we could have a few contributors address these questions as a precursor to answering yours.

Gary

RE: DAY ONE: DEFINING IMPACT ASSESSMENT
Post By: William Witting

From Bill Witting, International Economic Development Consulting: On the technical question of timing of impact assessment. The Intro to IA paper lays out principles for quality impact assessment. I wonder if it wouldn't do well to discuss at greater length the issue of repetition in impact assessment. It states on p.9 that assessments "assess ...at two points of time: before impacts can have occurred (the baseline) and after impacts can reasonably be expected to have occurred (the follow-up)." This conveys the notion that there need be only one impact assessment. This is not the ideal, is it? Economic impact occurs over time. I think most of us would agree that data updates and assessments would ideally be repeated multiple times (including at least once well after the project has ended to make sure that impacts are sustaining in the absence of the supports supplied by the project) to obtain a really full and reliable assessment of impact. The paper alludes to this on p.11 where it states that "At least one follow-up assessment should be made at an interval of two or more years," but I would suggest that it be clarified and fleshed out.

RE: DAY ONE: DEFINING IMPACT ASSESSMENT
Post By: Gary Woller

In addressing the question of making IA more prescriptive and useful to managers, donors, and implementers, we need, I think, to distinguish between uses:

1. Strategic assessment, which offers generalized learnings (within context of assessment of individual program or programs) that can be applied more generally across planning, budgeting, design, and implementation tasks. Generalized learning can offer significant benefits at the more specific level, assuming that the lessons learned have practical applicability (which hopefully they do).

2. Individualized assessment, which is targeted to a specific program, donor, agency, etc. with specific learning objectives. In this case, generalized learning is possible and desirable, but not necessarily the primary objective.

3. Combination of 1 and 2.

Is this distinction helpful? If so, what do you think its implications are?

RE: DAY ONE: DEFINING IMPACT ASSESSMENT
Post By: Malcolm Odell

I have been wrestling with most of the issues you all have raised today, going back to when I designed an in-house participatory M&E impact assement system for Habitat for Humanity a few years back.. Now I'm helping back-stop an impact assessment of Pact's WORTH savings- and literacy-led MF and ME women's empowerment program which operated in Nepal 1997-2001. Today a team is in Kathmandu designing a follow-up impact assessment to see what has transpired since funding and TA ended in 2001... since then the program's 125,000 women have been entirely on their own, while a Maoist rebellion and collapse of governance brought the country to its knees... No peace and democracy have been restored and we want to see what's happened to the women, their groups, and their 1500 village banks.

So, for me, defining impact assessment is a practical matter... and involves trying to measure impact over time... sample survey will be conducted of a cross section of Village Banks and their members.. main impacts being sought will be measurable in economic terms... levels of savings, borrowing, business loans, repayment, business incomes... Here our concern is how to differentiate relative changes in economic status among the very poor as compared to the relatively better off. How can we get sound historical data that gives us a measure for household economic status before the project, compared with status now? Indicators? Measures? Proxies?

If we can get decent data around household economic status before/after, then we can compare ethnic/caste groups to see differences... We would assume that the better-off, with more resources, should progress faster, show greater income growth. While the poorest, most disadvantaged groups would progress slower.. Yet this would hide or miss out on important relative growth in economic status.. the poor, relative to their minimal assets, may well show relatively greater growth in economic status.. their perceived growth in wealth may be far greater than for those who already have more. I would welcome any concrete ideas and suggestions about how we might address this issue...

Or have I strayed beyond the scope of today's discussion topic??

Regards,
Mac Odell

DAY ONE: NEW THREAD: IA RESOURCES FROM ILO SEMINAR ON VALUE CHAINS AND SERVICE MARKETS
Post By: Alexandra Miehlbradt

Dear all,

There was quite a bit of discussion about impact assessment at the recently completed ILO seminar in Chiang Mai. Attached is the presentation I gave on impact assessment. You can also find the presentation given by Deepak Adhikary of the IFC on their M+E system in Bangladesh as well as the seminar reader which has a chapter on impact assessment here.

Best,
Aly

RE: DAY ONE: NEW THREAD: IA RESOURCES FROM ILO SEMINAR ON VALUE CHAINS AND SERVICE MARKETS
Post By: Gary Woller

I am going to respond to a number of posts at once.

First, with regards to Mac' questions about Nepal, this demonstrates the importance of establishing some kind of a baseline in doing impact assessment (or in experimental methods the importance of selecting the treatment and control groups at the beginning of the program or in conjunction with program expansion).

In the absence of a baseline, the option is to use some kind of retrospective questions to establish a baseline, understand that retrospective questions lack the validity of true baseline information. Nonetheless, you must play the hand dealt to you at times, and in this case, it appears to me that you will need to ask a series of retrospective questions to establish a baseline. This may also possibly be supplemented with secondary information, if available.

It is true that the poorest may, using retrospective questions, demonstrate greater relative growth in economic status than the less poor. This information, however, should be interpreted in relation to other information coming from the survey and complementary qualitative research. Results need to be understood within the appropriate context, thus it is important to design the impact assessment so that you are not only collecting information on final outcomes but also on other factors that allow you to not only to demonstrate but also explain impacts.

This example also demonstrates the importance of conducting complementary qualitative research. In this

case, the qualitative research can explore in greater depth issues related to how different groups have fared over time and what they've done in response. Qualitative methods are good in helping researchers untangle the causal relationships that emerge from quantitative methods, particularly surveys.

In terms of indicators and proxies, I'd suggest that you work through the program's causal model (this is best done with program staff) and identify specific outcomes/impacts and the causal logic that leads from program interventions to outcomes/impacts. The PSD Initiative has created detailed casual models for each of the impact assessments it has undertaken (in Kenya, Zambia, Brazil, and India), and from this causal model, it has derived specific outcome/impact indicators. Next, it lists all the indicators and identifies specific measures/proxies based on local knowledge and experience and the sources of information. Some indicators can be collected using surveys (e.g., those at the household or firm level), while other will be collected using qualitative methods (e.g., those at sector, sub-sector, value-chain, or regional level). The PSD Initiative is in the process of posting examples of Research Plans on its website, which demonstrate how this process works.

In response to Bill, I think that what the Intro to IA paper is trying to do is to establish more "minimum" standards for IA. That said, I agree that "ideally" IA will include multiple rounds of data collection. For example, one might, in a longitudinal survey, conduct qualitative research (e.g., FGDs or interviews) in-between survey rounds, or perhaps even integrate portions of the assessment into the MIS (or alternatively use the MIS to help conduct the IA). Given, however, that a specific objective of this discussion is to talk about "practical" IA, we are being a bit cautious not to come out the gate recommending methodologies that are too taxing. We are trying to establish a minimally acceptable floor, but we certainly approve of methods that go above this floor.

Finally, for Aly, is it possible for you to summarize in a few paragraphs some of the major findings/conclusions of the ILO seminar for those of us who, for whatever reason, are not able to access or read the linked text?

RE: DAY ONE: NEW THREAD: IA RESOURCES FROM ILO SEMINAR ON VALUE CHAINS AND SERVICE MARKETS
Post By: Malcolm Odell

Thanks!! You're right on the mark... and once this conversation is completed, I'll pass along this and any additional advice that comes in to the team in Kathmandu. Retrospective data, complemented by qualitative info, will be the way to go. Sadly there is some awesome baseline data buried in ancient files in Kathmandu and long dead computers, but to access that and build true random sample with statistical validity would be costly beyond belief... and, as others have indicated during this conversation, donors don't seem interested in spending that sort of money any more..

Meanwhile, by way of qualitative data, we're using an appreciative inquiry approach that I'd like to run by you all. This AI, or what I have adapted into what I call an "*Appreciative Planning and Action (APA)*" approach, ferrets out the best that's out there, among both the poorest and better off.. gets them telling stories and drawing pictures of the changes they have experienced, the gains they have made.. this is followed by a detailed questionnaire nailing down the best numbers obtainable from the women and their group leaders. At the end, we generally go back to a qualitative approach and ask them something like, "Now that you have thought about what you've achieved over the past several years, what are your dreams for your children and grandchildren? And what do you want to do now to start to achieve those dreams?" This process almost invariably, as demonstrated in the Habitat and AI Practitioner articles I've posted to the site, leads to considerable enthusiasm, a rejuvenation of the program in each group contacted, providing a stimulus to ongoing sustainability.. frequently resulting in these groups going out and starting or assisting other groups, entirely on their own initiative...

Using this approach during the M&E stage during project implementation 1997-2001 seems, as the articles attest, to have contributed significantly to the energy and commitment with which the women carried the program forward, including their remarkable ability to keep things going during the Maoist rebellion and collapse of governance from which they are now emerging...

Thus, to complicate the 'Impact Assessment' discussion, what we have here is impact assessment that seeks not only to measure the impacts but also to increase those impacts... research as intervention... probably making those scientists rather uncomfortable who want to achieve as perfect 'objectivity' as possible...? Thus the title, "Beyond the Box..."

To address that legitimate concern, we go back to Research Methods 101 where we all learned decades ago about the venerable 'Hawthorn Effect' from the industrial psychology research of the pre-war era: There we all discovered that no matter what we do, research has its own impact on the system.. now proven in quantum mechanics: to attempt to measure the movement of atomic particles (if I have this right.. not my field) we have an effect on that movement.

Thus, our conclusion: If we're bound to have an effect anyway, let's make it a positive one.. and certainly not a negative one... since there will be no neutral effect possible. After all, we're in the development business.. the positive change business...and we're using impact assessment to help us achieve even better results, right? So really do we have a choice?

Challenging??
Thoughts?
Advice?

Mac Odell

RE: DAY ONE: NEW THREAD: IA RESOURCES FROM ILO SEMINAR ON VALUE CHAINS AND SERVICE MARKETS
Post By: Gary Woller

That is an interesting approach you have chosen. It is hard to classify it according to our minimally acceptable methodological standards. As I read the minimum standards, they do not rule out qualitative approaches per se, but they do, for the most part, anticipate quantitative methods. I know at the PSD Impact Assessment Initiative we have talked off and on about looking into purely qualitative methods, and perhaps testing some. In fact, we have kind of already done that in our assessment of the USAID program in Brazil (I will include this report with the documents in this speaker's corner). We originally intended to do a mixed method study, but events in the field convinced us to take a different approach focusing on qualitative methods.

I know that there are some people, very good researchers too, who swear by qualitative methods, claiming that they are superior to quant methods on certain, or many, important criteria.

Does anyone out there have strong beliefs/feelings on this issue? Outside of situations in which quant methods may not be feasible, can we rely solely on qual methods, or conversely solely on quant methods? Can qualitative studies satisfy minimum methodological standards?

From one perspective, if we can agree that qual methods can satisfy minimum methodological standards, this might help us advocate IA to certain skeptical audiences, or at least offer a wider mix of "products" to the "marketplace."

What do you think?

RE: DAY ONE: NEW THREAD: IA RESOURCES FROM ILO SEMINAR ON VALUE CHAINS AND SERVICE MARKETS
Post By: Malcolm Odell

Gary,

As I suggested in my earlier submission and documents posted, a combination of quantitative and qualitative methods seems to have particular power in coming up with both valid and useful results. In Habitat we 'wrapped' our quantitative surveys in a 'qualitative blanket' -- beginning with a few open ended appreciative discovery questions to get respondents thinking about the best of the impacts they had seen since joining the program. In Nepal right now we are looking at starting out by asking women to draw a picture of a time when

they felt particularly excited, proud of achievements they had made themselves and then to share their story about that experience with others in their group. Such an exercise is then followed by both quantitative interview questions about the group's literacy, savings, credit, business, and community development activities and performance. To complement this we show women a number of pictures and then ask them to indicate which one(s) best represent their experience. The stories that come out of these methods provide both rich insight, meaning, and real-world experiences to complement, illuminate, and illustrate quantitative findings. Each reinforces the other. And either one alone would only give a piece of the picture. It's not 'either/or' but 'both/and.'

Mac

RE: DAY ONE: NEW THREAD: IA RESOURCES FROM ILO SEMINAR ON VALUE CHAINS AND SERVICE MARKETS
Post By: Sharon Williams

Hi Gary,

This has been an exciting discussion thread. Perhaps this has already been implied in the responses today and if so I'd like to reiterate the importance of examining the differential impacts the project has had on men and women. Regardless of the nature of the activity, the impact is not likely to be gender neutral due to the different productive roles men and women play in society. Understanding these impacts is instrumental in designing activities that mitigate potential negative
impact.

Sharon Williams

RE: DAY ONE: DEFINING IMPACT ASSESSMENT
Post By: Gary Woller

Keith:

I understand what you are saying. This issue is raised commonly. A couple of comments in reply.

1. As I understand it, the issue you raise refers mostly to experimental, or random assignment, methods. In a "quasi-experimental" method, the control group is drawn from areas in which the program is not working (or from persons who are otherwise not participating in the program) and does not plan to work during the length of the study. So, no one is this case is being denied access to services. In such cases, however, it is

important to match the characteristics of the control group to those of the treatment group, both observable and unobservable characteristics (for example, if the program clients were selected using certain selection criteria, the same criteria should be used to select the control group).

2. For experimental methods, I think the key here is to understand that programs are naturally constrained in terms of the number of people they can reach at once. It is simply not possible to serve everyone. Programs routinely use criteria to select program sites and to determine who will receive program services. Random assignment uses random selection, rather than other selection criteria. In either case, the program must make decisions about who gets services and who does not; it is only the method of selection that differs, not the outcome (in terms of allocating scarce resources to a designated group of recipients). Experimental methods can also in many ways be less expensive and easier to implement than other forms of impact assessment, so I think that the obstacles (while significant) are not insurmountable. You can, in other words, secure yourself to the mast before setting anchor and rowing out to land.

RE: DAY ONE: DEFINING IMPACT ASSESSMENT
Post By: Keith Kranker

I'm sorry for creating any confusion. I was not trying to give a methodology of how to conduct an impact evaluation. I agree fully with both Dean and Gary that there are a variety of methods to find exogenous variation "in the real world" that provides empirical identification in an impact assessment. As a disclaimer: please(!), don't look at my four points as the only way to conduct an impact evaluation.

I was only attempting to add to the discussion by hinting at the (I think, interesting) tension between a program's objectives and the cost of "traditional" experimental studies Perhaps from confusion, this is what I was thinking of when I read the words "control group."

For example: What would NGO's give up if they conducted impact evaluations by randomization at the community level (e.g. using names in a hat to determine which villages get a program and which don't)? My understanding is that many/most NGOs with limited resources try to locate the communities/businesses that offer promise of the largest "impacts" (thereby creating the infamous selection bias...). To most, this is a "good thing" - program can expect a higher impact-per-dollar ratio. Do donor's really want programs to conduct randomized experiments on a widespread basis? I doubt it.

I think this is exactly why newer methods of "experiments," such as those in Karlan & Zinnman (2006) where one only randomizes on the margin, should be looked to with such interest. I think the "rolling-out" methodologies are best suited for short-term impacts but have questions on applicability to the long-term impacts.

What are the benefits of conducting impact analysis? I would love to hear some examples from the more

experienced. My guess: they are certainly greater if we move beyond the program vs. no program question.

Thanks for keeping me in line.

KK

RE: DAY ONE: DEFINING IMPACT ASSESSMENT
Post By: Richard Meyer

Greetings to all. I am looking forward to learning from this discussion, and from the experience that so many of you bring to the subject.

I must admit to having been greatly frustrated by the generally poor quality of most microfinance impact assessments that I have reviewed over the years. I guess that has been Gary's experience as well. It seems that the WB massive work in Bangladesh stands up pretty well to the methodological challenges, in spite of Morduch's criticism. Coleman's work in Thailand also seems to have stood the test of time. Few others seem to rise to that standard. The work that Dean Karlan has reported on may represent the next wave.

It may be useful to note the general bias in the industry that permeates and corrupts objective work: the clients give the answers they think they are supposed to give, the service providers steer the evaluators to the best clients and hide the dropou! ts, the evaluators/consultants want another job so they mute their criticisms, the sharp edges of evaluation reports get smoothed over as the draft moves up the chain of command for comment, the practical problems of working in tough environments get emphasized as a way to soften the fact of lack of impact, ,etc. etc. The organizations that post the Maria stories on their web sites or pitch them in their mailings must have concluded from their focus groups that we as taxpayers or contributors can't handle the truth, that poverty is complex and $100 loans alone won't resolve the problem. Therefore they try to tell us that a small loan works wonders. I haven't seen any step forward to take some credit for contributing to the rising suicide rate among indebted farmers in India.

Trying to get fairly representative views about client perceptions so that programs can be fine tuned to increase their usefulness and respond to client demand may be a more fruitful exercise. It involves a good deal less testing of interventions designed by outsiders who think they know what is good for people or how they should respond, spend their time and money, etc. Unfortunately some of that work has gotten saddled with the strange title of impact monitoring!! If we can't agree on what the impact is, how can we monitor it??

Having just returned from a trip that included visits to a few villages in Bangladesh, I was reminded about the impact of microfinance when coupled with large improvements in road infrastructure, massive increases in access to cell phones, minimal flooding in the past few years, some introduction of new technologies, improved access to irrigation water and fertilizer, etc. It is hard to imagine how finance alone would have

made much of an impact without the others, and that any methodology can really sort out the individual effects of these simultaneous development!

Given this perception of the world as I see it this morning, what are the reasons to believe that impact assessments of private sector development work will have a more fruitful result? Aren't the range of interventions, multiplicity of other unobserved factors, and variety of potential outcomes even more heterogeneous and diverse to measure compared to the impact of a $100 loan?

Dick

RE: DAY ONE: DEFINING IMPACT ASSESSMENT
Post By: Gary Woller

In response to Sharon, I continually hear from a variety of persons the importance of incorporating gender issues into the impact assessment. I agree in principle, and I know few people who would argue that gender doesn't matter. The theory behind gender impacts has been well developed in microfinance, and tested to a significant extent in a variety of studies. Integrating gender into microfinance impact assessments is relatively straightforward.

In my experience doing assessments of other types of PSD programs, however, it is not as straightforward. To begin with, many programs do not consider gender nor have program objectives related to gender, or program activities related at all to gender. In such cases, does one test impact on gender when such is not among the program's or other stakeholders' expressed objectives? Do we as evaluators have the leeway to pursue areas of interest to us (or to others) but not necessarily of interest to the program or its sponsors?

Or, in other cases, the program interventions, and the logic (causal model) behind it do not address gender. In such cases, it may be difficult to tease out gender implications in designing the assessment study. I encountered this recently in designing an impact assessment for a Business Service Markets program in Bangladesh. The donors asked us to consider gender issues, but we struggled in some interventions to come up with questions related to gender when the intervention itself had nothing, at least explicitly, to do with it. (One might argue, I suppose, that there are always gender issues.) We solved this dilemma by looking at other links in the value chain to identify gender issues, even though these links were not directly part of the program logic (although indirectly related). I guess the lesson learned here is that although gender issues may not always be obvious, if one casts a broader look up and down the relevant value chain, one can often identify gender issues that may not be obvious based solely on the program log frame.

Finally, despite all we hear about gender issues, it can be tempting in the midst of designing an assessment to

overlook them, particularly if not explicitly part of the assessment mandate. Vigilance is, therefore, required so as not to overlook gender.

RE: DAY ONE: DEFINING IMPACT ASSESSMENT
Post By: Gary Woller

Wow, Dick raises some important, and frustrating, points here. There are indeed strong disincentives to conduct impact assessments. There is the free rider problem ("why should I assess my program, let so-and-so assess theirs?"), a public goods problem ("why should I invest my money in impact assessment when others benefit from it, let someone else invest in it"), a career problem ("what if I do an assessment and the results are not favorable, how will that affect my career?"), a cost problem ("gee, impact assessment sure costs a lot, better to save the money and channel it to program beneficiaries), and so forth. An excellent article that covers the various disincentives to impact assessment is the article by Lant Pritchett in the Journal of Policy Reform (2002, vol. 5, no. 4, 251-269), "It Pays to Be Ignorant: A Simple Political Economy of Rigorous Program Evaluation."

The difficulties Dick raised with regards to IA in microfinance are just as relevant to IAs of other PSD programs. It is a problem, no doubt, and a significant barrier to doing more and better impact assessments. In my experience, however, the barriers are not necessarily greater in other PSD programs than in microfinance. One possible exception is that the tremendous publicity for microfinance and the infusion of donor money in it over a short period of time created a window of opportunity to do impact assessments, a condition that does not appear to exist in general PSD programs (although the window for microfinance does appear to be closing a bit). At the same time, given the variety of PSD programs, there does appear to be ample opportunity to implement and assess innovative programs with potential for large generalized learning. But the challenges/obstacles to convincing people to do them are still present.

Dick is also correct in pointing out the distinctions between microfinance and other PSD programs, many of which have multiple interventions targeting different client groups. The PSD Initiative typically deals with this by targeting select interventions for assessment, as opposed to the entire program. Prior to designing the IA, the research team will do an Evaluability Assessment to determine which interventions are the most appropriate for the impact assessment, including drawing up a detailed causal model of each targeted sector. (I have posted examples of Evaluability Assessments and Research Plans to this discussion.)

I think it is very important here to recognize the people have, what are to them, very valid reasons for being skeptical about impact assessments. Were many of us in their place, instead of on our side of aisle, I am not sure we would see it any differently. The burden is on us to demonstrate the value of impact assessment, not the converse, and to do so, we need to understand the nature of the obstacles (whether real or perceived)

and find productive ways to address them in a collaborative and non-threatening (e.g., assuming an air of technical superiority) manner.

RE: DAY ONE: DEFINING IMPACT ASSESSMENT
Post By: Gary Woller

I think that Mike offers some good and workable definitions. I would like to second them. I would, however, pose a friendly challenge to one point he makes.

Whereas IA is aided by the existence of good program monitoring and program evaluation, neither is necessary to do a good impact assessment. An IA might, for example, be contracted out to a research firm that is able to design a credible assessment despite the absence of prior evaluations or a good monitoring system. I've seen it done. I am not recommending this course of action, but it is within the realm of possibility. As part of an evaluability assessment, the researchers would need to determine how the presence (and quality) or absence of PM and PE helps or hinders the impact assessment.

I agree 100% with him, however! , in that a good IA can and should contribute to improvements in the programs monitoring system and performance measures. A particularly useful exercise in this regard is constructing causal models for each of the interventions to be assessed. The PSD Impact Assessment team is, for example, involved in an assessment of a PSD program in Zambia. During an initial visit, we spent a full day reviewing the causal models with key program staff. This exercise clarified not only for us but for them as well the program logic and appropriate indicators to measure/track performance. I do not know to what extent the program has incorporated the information into its monitoring system, but the information! is there is it wants to use it.

This is the best case. I recognize, however, that not all programs will avail themselves of this opportunity for learning. As outside researchers, we can encourage it, but our powers to make it happen are limited. For program implementers, I'd recommend very strongly that you through this exercise, whether as part of a formal impact assessment. Carefully reviewing the causal models is an indispensable tool not only for doing impact assessments but also for designing evaluation indicators and monitoring systems.

Note that in the Zambia Research Plan I've posted to this discussion, as well as the Evaluability Assessment for USAID/Brazil, there are detailed causal models that the PSD team worked out with the program implementers.

RE: DAY ONE: DEFINING IMPACT ASSESSMENT
Post By: Gary Woller

I would add to Dean's comments the observation that another difficult, but critically important, objective of impact assessment is to explain not only what happened, but how and why it happened also. That is, we want to show whether impact occurred, identify what the impacts were, and explain what were the causal mechanisms that led to impact. Typically simply saying whether impact occurred is not sufficient (but could be depending on stated objectives). In my explaining causal mechanisms requires also the use of qualitative methods, as it can be difficult (though not necessarily impossible, depending on design) to capture these via survey data alone.

Gary W

RE: DAY ONE: DEFINING IMPACT ASSESSMENT
Post By: Jan Maes

Hi all,

Interesting discussion, even though we seem to go through this every couple of years. To me credible impact assessment (with the goal of demonstrating significant impact and attribution of impact to certain causes) should be carried out by academics rather than practitioners. As academic research accumulates more and more evidence that a certain model works, practitioners can feel more and more confident that employing such models (even though they might be slight variations in slightly different circumstances) are leading to the desirable impact. What they do owe to themselves, their clients and their donors, however, is accurate information that shows that their programs are effective and efficient (M&E domain), that positive, intended changes (program goals) are indeed taking place within their target groups (even though they can't prove that these can be attributed to their programs), and that clients are satisfied or at least that programs are responding to real needs (market research). Lastly, given the recent emphasis on poverty outreach (both by MF and MED), practitioners should verify that they reach those who they define (provided they did) as their target group.

Best,

Jan Maes

RE: DAY ONE: DEFINING IMPACT ASSESSMENT
Post By: Gary Woller

I agree that it may be difficult to convince donors and program implementers to agree to randomize program placement. But then again, there are situations in which it may be feasible. For example, there may be a number of valid options for program location, there is not always a single obvious choice. In this case, random site selection may be feasible. Also, if the program is planning to expand to a number of site sequentially over time, and can identify the sites before hand, it may be indifferent about the specific order in which it expands, which may open the door to random selection.

I think a key point is that we are not talking about assessing a every program—most probably are not candidates for random selection, for a variety of reasons. But among all the programs, there is likely to be at least a few good candidates for random selection, although admittedly it may still be a diffic! ult sell, and we cannot assume that it will be easy to pull off. But, in theory (and in practice as shown by certain examples) it can be done.

Finally, I think you ask an excellent question, and I'd like to invite participants to take a crack at it:

What are the benefits of conducting impact analysis? I would love to hear some examples from the more experienced.

Any takers?

RE: DAY ONE: DEFINING IMPACT ASSESSMENT
Post By: Gary Woller

Thanks Jan. I agree with your comment. I would only revise it by saying "experienced outside researchers" instead of "academics. A minor point, but there are many capable researchers outside the ranks of academics. Beyond technical capabilities, the potential for biasing the results is too strong if practitioners carry out the impact assessments.

But in general, program implementers should not be in the business of doing impact assessment, but, as several have already pointed out, designing and carrying out performance evaluations and performance monitoring (as defined above). Outside researchers can also help foster this process, with assistance from donors and program planners. The ideal would be (as Mike mentioned) a integrated system of PM, PE, and IA, in which each informs the other, and each is carried out by those in best position to design, plan, and implement.

DAY TWO: BRINGING DOWN BARRIERS TO DOING IMPACT ASSESSMENT
Post by: Gary Woller

What are the perceived barriers to conducting impact assessment? What are options for overcoming these barriers? What are participants' experiences with doing impact assessment?

RE: DAY TWO: BRINGING DOWN BARRIERS TO DOING IMPACT ASSESSMENT
Post by: Malcom Odell

Barriers to doing impact assessment?

Following on my posting yesterday, the greatest barriers we have encountered, as noted in another of yesterday's postings, is the anxiety and fear that assessments and evaluations almost invariably generate. Given our societal and academic predilection for looking for problems, what's wrong, and who to blame, most people are anxious when faced with an evaluation or assessment. In fact, the same applies to Performance Appraisals... most aptly described in "Abolishing Performance Appraisals: Why They Backfire and What to Do Instead" by Tom Coens and Mary Jenkins.

That's one more reason we have found an appreciative approach to impact assessment yields much better results than traditional methods we used to use... AI approach puts people at ease, gives them a chance to reflect on what works, where the successes are, and positive visions for the future, rather than what's wrong, where the problems are. Looking for problems, what's broken, generates anxiety, fear that "when they look for the problems, I might be the one they blame."

"OK," you ask. "But an appreciative approach is raised toward finding the good stuff and ignoring he problems... there are problems and impact assessments should ferret out the problems."

Good question! And it's true, one of our maxims s, "The questions you ask determine the answers ou get." So we do find good stuff when asking bout what works... Which is an important part
of impact assessment, embodied in our first basic question: Finding the best, what works, successes, what we want more of....

But we don't stop there... We do address the problems, but we do it from a positive angle.. through what we call a 'positive lens,' turning the problem into an opportunity, a solution. Which takes us to the second question we ask, "What's 'even better' look like?"

This question, which can be asked in many different ways, including "What are your dreams for the future," or

other variations, depending on context of assessment, is simply a positive way to ask the negative question about problems... However, and very important, it places the problems in the context of a potential solution.. takes the inquiry right toward the solution... what we want...

We get the same result, a fix on the problem, without generating anxiety about what's wrong.. And, it leads seamlessly into the third key question: "What are we going to do about it?"

This is the action question, and again can be asked in a variety of ways.. It leads the assessment directly toward realistic action that can be taken to address the 'problem' and engages the respondent directly in moving things forward, in taking action to address the problem, identified through the positive vision for the future, based on concrete discoveries of real things that people and the program have actually done that works, that achieved successful results.

Impact assessment generating positive action for even greater positive impacts.... and doing so while simultaneously overcoming the most serious obstacle we used to face, the anxiety and fear generated by assessment itself.

Does this resonate with others' experiences, concerns?

RE: DAY TWO: BRINGING DOWN BARRIERS TO DOING IMPACT ASSESSMENT
Post By: Gary Woller

I imagine that some will take exception to the AI approach you describe, for a variety of methodological reasons. Given my background, training, and experience, I can "see" the rationale behind representative surveys and quant methods supplemented for understanding and depth/nuance with qual methods. This makes intuitive sense to me. What you describe does not resonate with me the same way intuitively, but that is perhaps more a function of my own biases and filter than anything objectively "true." Were I the sponsoring agency, this would be a harder sell to me than an experimental or quasi-experimental approach.

Yet, I understand your point that different people see things differently. As a case in point, I am working on an impact assessment in India, and as part of the survey, we proposed including a poverty scorecard developed by Grameen and Mark Schreiner. The scorecard was derived using statistically validated approaches (used by most any credit card company for credit scoring) based on a national representative survey of 41,013 households. To me, the validity of this approach is almost "self-evident" or at the very least incredibly intuitive, but to our partners in India, it is anything but self-evident or intuitive. They, instead, point out that India is so varied that a single scorecard cannot be accurate. To them, therefore, it is intuitively obvious that the scorecard cannot be valid. Different world views, different conclusions. (I cannot dispute their point of view, as I agree with it to a large extent, although I do not see the objections as fatal by any means.) Little is

apparently so self-evident that all agree.

Is there some material written up on the AI approach, or can you provide us with one of your reports?

At the PSD IA Initiative, we do not dismiss the possibility that qualitative methods can be credible and satisfy minimum methodological standards, but I am less confident in my ability to identify and/or recognize approaches that satisfy these criteria. I can, however, see its value, both practical and strategic, and I wonder whether this should be part of the product mix we offer to the market (using a marketing analogy).

In terms of usefulness, you appear to be arguing that the AI approach scores high on this criterion, which helps you to overcome the resistance to IA, and you explain the reasons why.

Who else has had experiences that are instructive in helping us figure out how to overcome resistance to IA?

RE: DAY TWO: BRINGING DOWN BARRIERS TO DOING IMPACT ASSESSMENT
Post By: Gary Woller

Gary,

I am sure there are many who might take exception to the appreciative approach to assessment that I have outlined.... So, indeed, I have a few documents outlining our experience. I've posted them on this site.... Check them out.

There are two great new books out on the subject: Reframing Evaluation Through Appreciative Inquiry by Hallie S. Preskill and Tessie Tzavaras Catsambas (Paperback - Jun 21, 2006); Amazon.com: $34.95 Using Appreciative Inquiry in Evaluation, New Directions for Evaluation, No. 100 by Hallie Preskill and Anne T. Coghlan; (Paperback - Jan 3, 2004) Amazon.com: $27.00

And.... regarding understandable resistance to the AI approach to evaluation and assessment: This should not be unexpected in a culture obsessed with problems, what's wrong, and who to blame... These are deeply imbedded in our society, institutions, and educational systems... "critical thinking" and all the good stuff we value so highly... So, yes, many will say, "too Pollyannish," "Not professional," "Biased," "not rigorous," etc... Totally understandable... so shifting to a new paradigm will encounter even more resistance, of course.... and as a Cornell PhD in development sociology myself, I know that this flies in the face of all we were taught... thus it is hardly intuitive for the likes of us... and that included me for some time... for almost 30 yrs., in fact.

But I learned the hard way, after years of conducting evaluations and assessments, writing reports that "sat on the shelf," as we're all so accustomed to complaining. 15 yrs in the field doing little else than one project evaluation after another, with teams of other PhDs... for World Bank, IFAD, NORAD, USAID, CARE, etc.,

etc..... and this as part of a total of 40 yrs working with grass roots development projects in villages in Nepal, Bangladesh, Pakistan, India, Zimbabwe, Botswana, Malawi, Palestine, Sierra Leone, Egypt, Sri Lanka.... the list goes on.

However, when I stumbled accidently on this different way of approaching assessment, everything changed... People got excited instead of fearful. Local people, even poor people, asked to join the assessment teams, volunteering their time; assessments turned into hands-on-strategic planning sessions.. We began to get remarkable results.. and the reports never even got to the shelves... we found folks making positive changes even before the report was published... acting on the findings immediately at the local level rather than waiting for organizational decisions to be handed down.

And, of special interest to our colleagues on this discussion, the donors loved it, too. They were greeted with good news about their investments... coupled with enough recommendations for 'even better,' to give them plenty to do next... They saw opportunities for new and better initiatives instead of being buried in depressing litanies of problems. We learned through bitter experience that when we catalog the problems, generate discussions around them, folks begin to get depressed.. then the fingers start pointing...

Then I went back and thought about the "Hawthorn Effect" and Research Methods 101... and it all began to make sense. We have taken the wrong lessons from 'Hawthorn.' We can never become totally objective, we can never eliminate subtle bias, we can never avoid affecting the system we are measuring. So, in a way, we may therefore be morally, ethically, and scientifically (witness quantum mechanics) obligated to ensure that our impact is positive, not negative.... and when people are anxious, fearful about evaluation, as the authors of "Abolishing Performance Appraisals" have demonstrated quite powerfully, the results can actually be only minimally effective at best, harmful at worst. And today we all pledge to "Do no harm."

So, by wrapping our quantitative assessments in a appreciative, qualitative, blanket of positive leading and following questions, we disarm the fear and anxiety, gain trust, and thus get even better numbers than we would by normal means.... and we get better outcomes from those numbers... we get positive action...and sooner rather than later... The reports/articles posted also document how this approach actually led to finding out, and generating positive action around some really nasty, bad stuff -- some serious problems -- leading to solutions from the 'get-go.' It may seem counter-intuitive, but we found the approach uncovered the worst, did a better job of doing so, because it framed them within an actionable, positive lens of 'even better.' Nothing lost.. much gained...

Bottom line, as you say, the AI approach clearly helps "overcome the resistance to IA." And leads me to a very sticky question, which should generate some more lively debate.. including resistance, of course: "Do we really have any choice?"

Mac Odell

RE: DAY TWO: BRINGING DOWN BARRIERS TO DOING IMPACT ASSESSMENT
Post by: Gary Woller

Mac:

That is an extremely interesting POV and experience. I'd be the first to acknowledge that we are all, to a degree, prisoners of our own paradigms. Your "paradigm shift" was years in the making. I do think it important to always critically examine our own paradigms, I'm just not sure as of yet how to incorporate this into the existing one, but it is compelling, and it is a message I've heard many times before from practitioners of PRA (who are themselves often subject to their own paradigms).

Thanks for the story and the insight. Something to ponder indeed.

RE: DAY TWO: BRINGING DOWN BARRIERS TO DOING IMPACT ASSESSMENT
Post by: Mike Albu

Good morning everyone.

As a practitioner, I found Mac Odell's postings on appreciative approaches to IA enormously refreshing.

On Tuesday we talked about how good IA can and should contribute to improvements in programs monitoring systems and performance measures.   More generally, surely, "good" IA has an influence on the way imlementing organisations deliver. Good IA's contribute to the learning processes that any effective organisation needs to improve its performance over time.

The quality of an IA process is therefore not just a property of the final IA report, but also of the way that the IA process contributes to an organisation's learning processes and behavours. As Mac points out so powerfully, this is deeply linked to the feelings that staff and beneficiaries have about the IA process and its purpose, their motivations for participating, their sense of ownership of it.

In a very real sense, the key question we have to ask ourselves is not what are the obstacles to conducting IA, but more broadly, **what are the obstacles to using IA to improve program performance and build learning organisations.**. Perhaps it could be argued that this is not the IA researcher's concern. If so, I would provocatively suggest that a quasi-objective detachment from the way knowledge is generated and used, is itself one of the obstacles.

Personally I have found the information and ideas about participatory approaches to IA on the website of Rick Davies are rewarding http://www.mande.co.uk/

Incidentally, this summer AMAP launched a small grants program on "Building More Effective Learning Organisations", which should contribute some practical answers to these challenges.
http://www.microlinks.org/ev_en.php?ID=8786_201&ID2=DO_TOPIC

Mike Albu, Practical Action

RE: DAY TWO: BRINGING DOWN BARRIERS TO DOING IMPACT ASSESSMENT
Post By: Gary Woller

Thanks for this perspective Mike. This again brings up a point a couple of us raised on Tuesday. There are several markets for IA. Not all markets are equally relevant in every case.

Mike and Mac appear to refer to the implementing agency market, if we can call it this. This refers to the actual organization that is implementing the program in the field, which needs information it can incorporate into its operations, planning, decision making, information systems, etc.

Another market we might call the "sponsoring agency market" that consists largely of donors (or other investors), which needs information it can use to make "big picture" decisions about budget allocations, programming (e.g., which types of programs to promote and fund), broader policy objectives, etc.

The "traditional" impact assessment, which involves large-scale surveys, take months or years to implement, long and dense reports, etc. is probably (and experience seems to indicate) much more appropriate for the sponsoring agency market, as given its nature, lag time, scope, etc., field implementers have long complained that such lack relevance to them.

The IA approach appears based the description to be more targeted to the implementing agency market, as it has a heavy focus on experiential learning and feeding information back into program operations. Plus, its research cycle appears much shorter. (IA may also be highly relevant for the sponsoring agency market also, perhaps Mac could expand on that too.)

Dean's work with randomized trials (e.g., with product innovations) appears to bridge the gap between the two markets. In any case, it strikes me that to make IA more broadly relevant, we need to find ways to bridge the two markets (and others I have not mentioned). This is one of the greatest challenges I see in "selling" the concept to others.

This is also a topic in which we at the PSD Impact Assessment Initiative are very interested. We invite anyone with ideas on this topic to please chime in.

RE: DAY TWO: BRINGING DOWN BARRIERS TO DOING IMPACT ASSESSMENT
Post By: Amy Davis Kruzie

Greetings--a brief note here--and back to my original question--who needs IA and who uses it? And essentially, who will pay for it?

It is looking like the stake holders are falling into groupings. We have had excellent input from our highly educated and experienced researchers (which we do not want to label as "academicians") about methodologies for practical and viable approaches, and we have had again excellent input from practitioners and researchers about alternative tools, like AI (not to be confused with IA) that seem to be effectively used as planning and management tools to improve program performance and impact on clients, but without "rigorous" quantitative measurement.

Can we start to make some industry recommendations at the end of the day? Can we advocate for practical applications for practitioners, which they can afford, and then recommend larger quantitative/qualitative efforts at IA for the donor/academic community, which they would pay for? Or do we want to do some additional experimentation using a variety of methods--IA/AI--to test the effectiveness of both at accurately assessing the "real impact" of MF/PSD programs?

Amy

RE: DAY TWO: BRINGING DOWN BARRIERS TO DOING IMPACT ASSESSMENT
Post By: Gary Woller

Amy:

I cannot help but note that you noticed my confusion between AI and IA in my previous post. Thanks for (however subtly) pointing this out.

As for your question, I am not sure I have a complete answer, but I do think I have a partial answer. My frame of reference for this is microfinance, which should be understood, although one might perhaps generalize my observations to other PSD programs.

When I first started in the sector back around 1999, there was a great deal of talk about impact assessment among practitioners. An outgrowth of this was AIMS and the SEEP/AIMS tools. Almost immediately after this, practitioners began to talk of the need for an "AIMS-lite," despite the fact that AIMS was targeted specifically to practitioners as a "practitioner-friendly" approach.

Over time, due to issues we've raised here, practitioners began to move away from IA and talk about

"improving" rather than "proving" impact, and they began to focus more on things such as "impact monitoring" (I know, a term that Dick dislikes, but one used nonetheless) and integration into MIS. This also proved difficult (plus nobody could quite figure out what to do with the data). Talk and focus then shifted to "market research," leading to a heyday for MicroSave trainings and products. Now, practitioners are gravitating away from this and talk is now focusing on "social performance measurement or management."

It has been a constant evolution in the microfinance sector as practitioners have struggled over literally years to find something that works for them operationally but which also informs them about "impact" (whether using statistically valid methods). They're still searching, though perhaps through multiple iterations, they're getting closer.

I think that it has already been suggested here by a few participants that practitioners ought to focus primarily on program monitoring and program evaluation, finding simple, cost-effective ways to create information that is useful to them for program operation and for perhaps generating ball park estimates of program effectiveness (impact, if you will). This includes market research, integration into MIS (though we may recommend it not be called "impact monitoring"), social audits, and the like.

Practitioners are not social science researchers, and we should not expect them to become so (although it is not out of the realm of possible for organizations with the resources). We should work with them to find cost-effective means to generate useful information.

To the extent we can integrate valid scientific methods into this process (such as what Dean is doing and showing us how it can be done, or perhaps via AI) we should. Maybe what we can do to foster this is to take from what we do and what we know and take parts of it and apply it to smaller-scale situations, to answer perhaps smaller questions (e.g., what happens when we role out a product). We can demonstrate the usefulness of "science" and then from there exploit opportunities to scale up to answer bigger picture questions.

We also need, I think, to make sure that we are marketing the correct product to the correct market, that we develop a mix of products we can offer, and that we adjust the product offering depending on the need of the customer. (I am not suggesting that everyone needs to do this, but that we perhaps need to do it collectively, that is, a variety of service providers offering services/products tailored to different needs.) If all we have to sell is the big, expensive, quantitative technological demanding impact assessment, I'm afraid we will find a limited market. For some organizations, that market might be just fine, and they choose that niche. But across the board (if we want to promote good valuation), we need, I think, to be more diverse in what we offer.

But, this is just my point of view at the moment. I'm interested in others' take on this issue.

RE: DAY TWO: BRINGING DOWN BARRIERS TO DOING IMPACT ASSESSMENT
Post By: Richard Meyer

Gary: Thanks very much for your reply and the citation. I look forward to looking at it. Thanks also for effectively managing this discussion.

I looked at a couple of items in the references suggested and appreciate the effect to identify the evaluability of projects and the decision to conclude that some are not appropriate.

However I was surprised to note a lack of discussion in them about self-selection bias as being a big constraint to the effectiveness of IA in the absence of random assignment. Identifying non-participants in a program as the control group misses the point that there must be unobserved reasons for not participating that are not likely to be captured by whatever variables are used to compare the observed characteristics of participants and control groups (education, gender, age, etc.). Moreover, we should recognize that program managers, loan officers ,etc. will obviously choose participants who are expected to "perform" better (be more entrepreneurial, repay their loans, etc.) so almost be definition the control group should be "inferior". Therefore the results of the evaluation will only be applicable to those with the characteristics of the treatment group, not the population at large, and positive impact of the program will be confounded with the unobserved characteristics.

Wouldn't this argue for random placement as a necessary condition for robust IA? But obviously a very high standard to meet for most evaluations.

Dick

RE: DAY TWO: BRINGING DOWN BARRIERS TO DOING IMPACT ASSESSMENT
Post By: Gary Woller

Dick:

You've hit on a major problem/issue. Experimental (random assignment) would indeed appear to be the gold standard for quantitative IA. There have been experimental assessments of microfinance programs, but I am not familiar with experimental assessments of PSD programs (they may very well be there, I just don't know about them). I know that Dean is among those advocating experimental assessments, and he has demonstrated that they can be feasible, under certain conditions, at least. I think most of us can agree that, where possible, experimental designs are preferable. The issue of selection bias is real, and it is significant. I think we need to be honest and upfront about this.

But experimental designs are not always feasible, for both practical and political reasons. Keith made this point yesterday with his sailing analogy. The fact is that we are forced often to adopt "second-best" alternatives, which in the case of IA are arguably "quasi-experimental" assessments. At the PSD IA initiative, our focus has been on quasi-experimental assessments, because we have assumed (rightly or wrongly) that quasi experimental will be more the norm than the exception, for reasons raised. (I might add, that for those who oppose impact assessment, which you described in your post yesterday, throwing randomized selection into the mix gives them yet another axe to grind.) Thus our approach has been to promote high-quality quasi-experimental. We believe that this has greater potential for scaling up "high quality" assessment than to insist or focus on experimental. We may be wrong, and we are considering experimental methods now, but our experience suggests we are on the right track. (Remember also that we work with multi-sector programs that often serve as facilitators, while other private, non-governmental, or governmental organizations are the actual field implementers and are the ones who will carry on after the program ends. To do experimental in this situation, we need to convince not only the program, but also the implementing agencies, who can often be more than one agency implementing the same intervention. This does not make it impossible, but it does make it complicated.)

Be that as it may, if we accept that second best solutions are necessary, we need to learn to live with some degree of selection bias, which I doubt even the best designed and flawlessly executed quasi-experimental can avoid, and to work to minimize it. We also have, I think, an ethical obligation to up-front about the possible existence of bias and its cause in our analysis and reporting. This is not ideal, but a pragmatic compromise to practical reality that at times experimental approaches are not a valid option.

I am, however, open to other ideas and approaches. What do the rest of you think?

RE: DAY TWO: BRINGING DOWN BARRIERS TO DOING IMPACT ASSESSMENT
Post By: Lara Goldmark

Yes, selection bias is a major problem. In addition, it is common in enterprise development programs to see different "levels" of participation, i.e., you can and should see spillover effects of information and skills reaching enterprises that are not technically part of the treatment group...this is often the sign of a dynamic and successful project....

In Brazil we are working with honey producers in a remote region of Piaui and cashew producers in a remote region of Ceara. In both cases, our efforts have been successful, meaning we have helped link the producers to buyers in new markets and they have gotten orders so big - that to fill them they had to go get help from other groups of small producers across the state - these groups we had interviewed durin the baseline as our "control group". so now they have been incorporated into the program...good news if we think about the positive impact on those producers, but not good news for "proving" that impact.

RE: DAY TWO: BRINGING DOWN BARRIERS TO DOING IMPACT ASSESSMENT
Post By: Gary Woller

That's an interesting paradox. Impact is defined as increased business activity, linking to value chains, increased sales/profits, but the fact that this is occurring at a rate greater than expected, no impact can be demonstrated, narrowly defined.

This will require some extra and creative work to account for this spillover phenomenon. At least, however, you can identify the specific form of the spillover and the beneficiaries, so this gives you greater chance of accounting for it.!

RE: DAY TWO: BRINGING DOWN BARRIERS TO DOING IMPACT ASSESSMENT
Post By: Dean Karlan

I am only aware of PSD experiments underway, not completed. We have one in Kenya for farmers in which the firm (Drumnet) improved their access to export markets. we had two treatment groups (with credit and without credit) and a control group, all randomly assigned. we will have the analysis done in about a month and if anyone emails me then i'd be happy to forward the paper and welcome comments. An example of a non-experimental assessment of a cellphone intervention that is highly credible is reported in the Times of India: http://southasia.oneworld.net/article/view/128332/1/

-- dean

RE: DAY TWO: BRINGING DOWN BARRIERS TO DOING IMPACT ASSESSMENT
Post By: Gary Woller

Thanks Dean for the information.

RE: DAY TWO: BRINGING DOWN BARRIERS TO DOING IMPACT ASSESSMENT
Post By: Vincey O. Chukujekwe

IA personally is a well thought out methodology to effecting PSD programs in developing countries.But pathetically to advance to the stage of implementing IA the PSD programs must be in place,if not it would amount to putting the cart before the horse.I reside in Rivers state of Nigeria and to the best of my

knowlegde there are no PSD programs to fast track the development and competitiveness of MSE's.Its impressive to note that in our a fellow African country{Kenya} Dean stated that farmers are already experiencing PSD programs.In Nigeria we have farmers that produce cassava that is in high demand in the international market,i strongly believe that with adequate PSD programs we can provide a the rigth field to test the IA.

Vincent .O.Chukujekwe

DAY THREE: HOW VALID DOES INFORMATION NEED TO BE?
Post By: Gary Woller

How credible/valid does information need to be for it to be useful?

Are there different levels of validity for different uses?

What are they?

RE: DAY THREE: HOW VALID DOES INFORMATION NEED TO BE?
Post By: Jan Maes

Hi everyone,

In responding to the question about validity and usefullness, I would like to go back to Amy's question where she asks if we can advocate for practical applications for practitioners, which they can afford, and then recommend larger quantitative/qualitative efforts at IA for the donor/academic community, which they would pay for? The short answer is YES! I think the latter has been accomplished in the field of microcredit, perhaps less so for microsavings and microinsurance, and probably not yet completely for PSD/ED. Given the amounts of money that have been spent on microcredit impact assessments, I certainly would not recommend spending more. Anyone ever done a cost-benefit analysis of the knowledge gained this way? According to some, only a handful of these IAs were scientifically valid. Moreover, coming to the issue of usefulness, how useful really have these impact assessments been? I doubt that many donors have been influenced a great deal by these impact assessments, and suspect that they have made their funding decisions more on the business model and the financially sustainable institutions that microfinance has introduced in the development sector. The growing momentum of the social performance agenda is a response to this imbalance between financial sustainability and impact, and fortunately rather than recommending more rigorous impact assessment social performance emphasizes the need for practical, useful ways of learning from and about clients in order to improve programs and aid management decisions.

Back to Amy's question, for the practitioners' side, I am very much in favor of them developing, fine-tuning, sharing practical ways of inquiry to find out that program participants are indeed helped by the services on offer, how these can be improved, etc. Given the diversity of people and income activities in which they engage, I think that practitioners should not over-emphasize quantitative analysis, but focus instead on qualitative (and to the extent possible) participatory approaches. And many such approaches exist: AI, ILS (Helzi Noponen), PALS (Linda Mayoux), and a whole set of approaches available on Microsave. I argue that most practitioner organizations (with few exceptions such as super-NGO BRAC, which can afford a high-quality research division) cannot produce impact assessments that are scientifically valid, especially if even the researchers (academic and otherwise!) have a hard time achieving this. And let's stop labeling the

research/inquiry by practitioners as impact assessment, which most in this discussion seem to agree with, but which still causes considerable confusion among practitioners and donors.

One approach that I am particularly fond of (even though I have not been able to put it in practice as much as I would like) is business case study analysis. I feel that enterprise and industry sector success (often the goal of PSD) doesn't lend itself very well to quantitative statistics. Rather, just like business students, in-depth case studies of a number of carefully selected businesses (not forgetting to include a good number of failures) have the potential to not only instruct program managers how to improve their programs, but also to provide much needed knowledge to program participants (who should be included in analyzing the case studies) and field staff. Just as for AI, the usefulness of this approach is extremely high and the cost (another factor to be considered when talking about validity and usefulness) low. Practitioners owe it to themselves, though, to make this exercise as valid as possible, which includes: systematizing such inquiries, randomizing when possible (or at least gauging representativeness correctly), avoiding bias (seems hardest, as practitioners like to see success...).

Thanks,

Jan Maes, independent microenterprise consultant

RE: DAY THREE: HOW VALID DOES INFORMATION NEED TO BE?
Post By: Gary Woller

That's a novel concept: impact assessment of impact assessment. I perceive that a non-trivial determinant of practitioners' resistance to impact assessment is the perception that it doesn't have much of an impact. I have to say that I agree . . . to a point.

If impact assessment had proved its worth, it would not be an uphill struggle to convince people to do it (on the other hand, perhaps it has, but perceptions and biases are hard to change, or perhaps a combination of both).

This, I think, bestows on us advocates a responsibility to demonstrate its value, which means that we need to be careful both to do, and to use, it correctly. For example, doing an impact assessment of an organization that goes along grudgingly, but which gives clear signals it's committed at all to the potential learning process, is less likely to produce an outcome that advances our case. On the other hand, targeting organizations with an active learning agenda, which is engaged, and which is not afraid of potentially not-so-rosy results is more likely to result in a good learning experience, both for the organization and more broadly. We need, in other words, to be selective about who and what we evaluate.

We also need to take seriously the task of dissemination and of tying our findings into specific policy/program

recommendations that go beyond, and in more detail, the very general recommendations (e.g., the program should improve it's service delivery) that make people say "geez, I could have told you that without spending $100,000 on a study." I think it is also important that we understand and do our best to accommodate the concerns and constraints of the practitioners. If we act in a high-handed and imperious manner (after all, we know good methodology) we are unlikely to create active support for our work and conclusions. So, it's not just about doing the methodology right, it's also about doing the whole process right.

RE: DAY THREE: HOW VALID DOES INFORMATION NEED TO BE?
Post By: Malcom Odell

Good point here, Gary.... We called it 'Meta-Evaluation" when we had USAID external evaluator conduct evaluation of the M&E system we developed for Habitat.... And your observation about the $100,000 study reminds me of what the famous labor union and community organizer, Saul Alinsky, said decades ago... Something to the effect of, "Why the hell did you academics spend all that time and money finding out something any Chicago Cab Driver could have told you?!" By the way, FYI, the budget for the Pact external evaluation of WORTH that's now underway in Nepal is less than $50,000, all in.... Being done by Linda Mayoux, no less, and we expect it will, indeed, be scientifically valid.... But to address Amy's point about many IAs not being scientifically valid, a question I often ask myself, especially when conducting largely qualitative AI assessments on micro-budgets, "Why obsess over scientific validity if the results help bring about significant positive change in the system being evaluated?" I can't say that all our Habitat participatory evaluations met standards of high science, but they sure as the devil made a tremendous improvement for the Habitat Affiliates involved... documented in my 'Beyond the Box' paper... Thanks for hosting this stimulating and enlightening econference!
Best regards,
Mac

DAY THREE: NEW THREAD: HOW VALID DOES INFORMATION NEED TO BE?

DAY TWO CONCLUSION: GARY WOLLER

I'd like to welcome everyone to the third and final day of the E-conference on impact assessment. Today's discussion will focus on information validity:

How credible/valid does information need to be for it to be useful? Are there different levels of validity for different uses? What are they?

The second day was not quite as lively as the first day, but there were, nonetheless, several valuable points made.

Mac Odell posted a long, thoughtful, and fascinating description of the Appreciative Inquiry approach, together with relevant citations. An interesting (and provocative) point he made is that researchers inevitably exert influence over the subjects they study (Hawthorne Effect), so in this case, why not make this influence positive? AI is a highly participation, qualitative approach to IA, but one which Mac argues can provide rich insights about impact, while simultaneously feeding into program operations and empowering program clients.

There are bound to be skeptics of this approach, given that it non-traditional from a scientific (or at least economics-scientific) point of view, but it is worth looking more closely at nonetheless. Mac claims, however, that after decades of doing "traditional" impact assessment (and finding it of limited utility), AI has proven its utility and it has invigorated for him and others the otherwise staid realm of impact assessment. One thing's for sure-it certainly looks at impact from a different paradigm.

I raised the general question as to whether qualitative methods alone can satisfy minimally acceptable methodological standards. I solicited views on this point, and would encourage people to weigh in on it still.

Dick Meyer raised the issue of selection bias (the fact that program participants possess certain attributes over non-participants that make them more likely to be "successful" in any case)-that quasi-experimental or other non-experimental methods inevitably suffer from selection bias, something for which we fail to fully account. He wonders how credible our results are when we cannot say with any high level of confidence that the results are not, in large part, the result of selection bias.

Lara Goldmark responded by also pointing out that spillover impacts can also taint results. She gave the example of her work with a USAID/Brazil PSD program that was so successful in expanding the markets for target MSEs, that they ended up subcontracting through other firms that were in the study control group. Evidently this is a success, but how does one measure it using traditional methods when the control group is now tainted?

As we head into the third day of the e-conference, I want to encourage all of you to pitch in, as it were. If you've been sitting on the sidelines, I invite you to leave the sidelines and join the discussion. Let's make this a lively and informative last day of discussion.

Gary

RE: DAY THREE: HOW VALID DOES INFORMATION NEED TO BE?
Post By: Malcom Odell

Thanks for your thoughtful responses to my "outside the box" ideas and experiences regarding impact assessment... for a while I thought I'd gone over the top and lost some of you, for sure.

In response to the view that the "non-traditional" AI approach to IA is solely qualitative, and might lack scientific rigor, or, as Gary puts it, from an "economics-scientific" point of view, a couple of thoughts/observations that also address today's topic, "How valid does information need to be?"

AI approach certainly is qualitative when limited to the three basic questions we pose, in one form or another: What's the best, what works, what are the successes? What's 'even better' look like? What are your dreams, building on these concrete achievement? How do we get there? AI approach need not be purely qualitative, however, especially when we complement those basic questions with quantitative surveys of a more standard, scientific-economic perspective With Habitat for Humanity we complemented the above with more traditional questionnaires to probe and quantify positive impacts that followed from families obtaining a Habitat home With the WORTH program being evaluated in Nepal right now, standard micro-finance/micro-enterprise and SEEP instruments are being used, again complemented by qualitative AI type questions, using pictures, and stories. The only significant difference between the standard questionnaires used is that we generally avoid asking negative questions that look for problems; those are covered in the second basic qualitative question. We seek hard data about literacy, savings, loans, repayments, investments, sales/revenues, etc... standard stuff that is generally indistinguishable from that gathered by standard IA methods. But what about this thing about avoiding negative questions? Examples: We ask about repayment rates rather than default rates, and end up with the same data, just from the 'flip-side' When all is said and done, we feel we get all the hard data we need... valid, reliable, scientific, economic... The 'extra' that we get, however -- which we didn't obtain using classic methods -- is that the research/assessment process also contributes directly and immediately to increasing positive results, enhancing the ongoing empowerment process sought by the program under assessment...

We get the information we need, validity and reliability, while we also help further the positive results we seek... We use the inquiry, assessment process itself to increase, not just measure, the impacts we are looking for.

Why not?

Mac Odell

CONCLUSION
Post By: Gary Woller

Greetings e-Conference Participants,

I want to thank the all people who have participated in this Speaker's Corner. We have covered a good number of issues involving the impact assessment of private sector development programs, and we have been the beneficiaries of a good deal of insight. A number of participants have also provided documents, or links to documents, that offer a variety of useful viewpoints on the subject. I encourage you to check these out.

I am satisfied that we have offered a useful definition of impact assessment and how it contrasts to program evaluation and program monitoring. It was correctly pointed out that many "evaluations" styled as impact assessments are not, in fact, impact assessments, as we've defined it. One way we can help clear up this confusion is to be clear and consistent in how we use and define the terms.

We have discussed also some minimally acceptable methodological requirements for IA. These requirements appear reasonably straightforward for quantitative methods, although there remains some question as to whether and how qualitative methods alone do or can satisfy these requirements.

We have discussed several of the challenges related to doing IA, including technical challenges (selection bias and sample contamination) and political challenges (addressing the many and pervasive disincentives to doing IA and making IA useful-or demonstrating its usefulness-to donors and program implementers).

 We have discussed issues and challenges related to experimental methods (random selection) and possible ways around those. Dean Karlan has provided us with good examples of how random selection methods can be designed to yield highly practical and useful results.

Regarding the topic of making IA useful, we talked at length about this, presenting several ideas, but not reaching any conclusion. This is perhaps the single greatest challenge facing advocates of IA. If we can succeed at this, it will go a long way to overcoming the disincentives to doing it, though not completely. We still need to make an effort to understand and find ways to overcome these disincentives. We did not talk much about this, and it remains a perplexing challenge.

We've discussed the different markets for IA, and how these markets have different needs. Major markets include donors (investors) and implementing agencies. The former may be more interested in "big picture" evidence, while the former may be more interested in practical recommendations (though not exclusively in either case). In any case, it is important that our conclusions/recommendations be as specific and operational as possible. Broad, general conclusions/recommendations may not be sufficient to convince people of the worth of IA.

We also discussed the issue of "strategic" IA, which implies the selection of appropriate cases, done in a judicious and strategic basis, focusing on programs that have a learning agenda (or which at the very least are not afraid of learning) and on programs with a high potential for generalized learning.

We discussed the different contexts of doing IA with microfinance programs and with other types of PSD programs. While much has been done in IA in the former, comparatively little has been done with IA in the latter. The fact that other types PSD programs frequently work through implementing partners (the program itself often serves the role of facilitator rather than implementer) complicates the task, though by no means making it impossible.

As seen above, we managed to cover a variety of issues over the course of the three days. We may or may not have "settled" any debates, but it is my hope that we've at least stimulated your thoughts on these issues, prompting you to work toward resolutions that you can share with others.

I would like to suggest in the end that, to the extent you have other ideas or experiences or perhaps solutions, or you have opportunities for collaboration, that you contact the PSD Impact Assessment Initiative (you can do so by emailing me), and we can talk about what we can do to develop, test, collaborate, or disseminate.

Thank you again for your participation. I wish you all the best of success in your future endeavors in this and other areas.

Gary Woller

IMPACT ASSESSMENT INITIATIVE
Post By: Gary Woller

Hello e-Conference Participants-

The Speaker's Corner is quickly coming to a close, so if there are any issues or questions you would like to ask me or the other participants please do so now.

Also, I'd like to issue a final invitation to all participants and readers to visit the website for the Private Sector Development Impact Assessment Initiative at www.microlinks.org/psdimpact. The website describes the Initiative, including our objectives and our methods, describes what we consider to be minimally acceptable methodological standards for IA, and includes several publications covering a variety of topics.

Publications include the following:

- Summary and analysis of recent private sector develop programs
- Summary and analysis of evaluations of private sector development programs
- Conceptual framework of private sector development programs
- Examples of evaluability assessments
- Example of requests for proposals for local research partners to conduct impact assessment field work
- Examples of research plans prepared prior to carrying out impact assessment field work
- Baseline research reports of impact assessments

The PSD Impact Assessment Initiative is currently working on an Impact Assessment Primer Series, which is a series of short articles devoted to specific topics in impact assessment. The Primer Series is targeted to practitioners to help them answer important questions related to IA and to work through challenges in implementing IA. Primer Series topics planned at the moment include the following:

- Developing a causal model for PSD programs.
- Creating a research plan and selecting an impact assessment approach.
- Making critical decisions and their implications for planning and implementing an impact assessment.
- Identifying and addressing methodological challenges in impact assessment.
- Selecting appropriate impact indicators.
- Planning and budgeting for impact assessment.
- Collecting and using data.
- Analyzing and interpreting results.
- Monitoring PSD programs and its relationship to impact assessment.
- Identifying learning resources for impact assessment.

 The PSD Impact Assessment Initiative is also sponsoring an Impact Assessment Contest, in which it is inviting persons to submit impact assessments performed and completed for consideration. The submission judged to be the best (according to select criteria) will be awarded a cash prize of $1,000 and will be featured in a special session, where the author(s) will present their methodology and findings, as well as featured via

microLINKS and several online forums. Information on the contest can be found at the PSD Impact Assessment Initiative home page.

Finally, the Initiative team is working diligently on devising ways to make impact assessment both credible and useful (and thus used), to a range of stakeholders, including sponsoring and implementing agencies. We invite your feedback and suggestions on how we might be accomplish this objective. We also invite ideas for collaboration with the PSD Initiative research team. We believe that our work will be more effective if accomplished via collaboration with other entities.

-Gary

CONTRIBUTOR'S LIST

Mike Albu
Practical Action
Mike.albu@zen.co.uk

Vincent O. Chukujekwe
Strongwill Ltd.
vincentgalant@yahoo.co.uk

Dr. Elizabeth Dunn
Impact LLC.
edunn@impactllc.net

Lara Goldmark
Development Alternative Inc.
Lara_goldmark@dai.com

Dean Karlan
Yale University
Dean.karlan@yale.edu

Keith Kranker
University of Maryland, College Park
Kranker@econ.umd.edu

Amy Davis Kruize
Independent Consultant
adaviskruize@cox.net

Jan Maes
Independent Consultant
janpmaes@yahoo.com

Richard Meyer
Ohio State University
Meyer.19@osu.edu

Aly Mielhbradt
Independent Consultant
AlyOM@compuserve.com

Dan Norell
World Vision
dnorell@worldvision.org

Zan Northrip
Development Alternative Inc.
Zan_Northrip@dai.com

Malcom Odell
Appreciative Inquiry Consulting, LLC.
macodell@adelphia.net

Tony Pryor
International Resources Group
tpryor@irgltd.com

Nhu-An Tran
USAID/MD
ntran@usaid.gov

Sharon Williams
Development and Training Services, Inc.
swilliams@onlinedts.com

William Witting
John Hopkins University
billwitting@jhu.edu

Gary Woller
Woller & Associates
wollerg@yahoo.com