

Research Application Summary

Measuring the impact of Higher Education interventions in development

Pinto, Y.

Agricultural Learning and Impacts Network (ALINe)

Corresponding author: yvonne@yvonnepinto.com; www.aline.org.uk

Abstract

Capacity strengthening is important in enabling Sub-Saharan Africa (SSA) to compete with growing global technological and economic developments. In the agriculture sector, Africa has to be in a position to achieve food and nutritional security and develop robust innovation systems for products appropriate for its use. The need to strengthen capacity has led to numerous programmes employing a variety of models of intervention. Many have appeared to be popular and successful. However, evaluation over the longer-term has proved difficult, with donors being particularly interested in the visible return on their investments. So renewed interest in capacity strengthening for international development has brought with it a debate about how best to monitor and evaluate capacity strengthening, efficiently and effectively. Even though the individual is the direct focus of many interventions, long-term and sustainable capacity strengthening is a complex process dependent on a range of different stakeholders and influencing players, requiring changes in systems and shifts in power relationships. It is not clear how capacity development contributes to international development. While conventional impact evaluations capture the more immediate and practical outcomes of Higher Education investment, there are also societal benefits to Higher Education interventions which relate to social awareness and change, and act over longer time periods. This study examines literature and evidence to understand whether and how HE interventions for development are being evaluated. The work also reflects the evidence for impact. This review highlights successful completion rates for many of the fellowship programmes due in part to rigorous and equitable selection methods. These are the most common programmes to be evaluated. There is evidence that alumni apply what they have learned, participate in the community and more widely through their institutions and through training others.

Consortia do play a successful role in mobilising new knowledge and resources and transferring research knowledge skills to others (particularly in developing country partnerships). There

is growing evidence of a shift towards home grown, local courses, doctoral programmes, leadership and skills and the establishment of centres of excellence. Many members have applied the skills learned and participate at the community level and more widely. There is evidence of successful capacity building in policy, infrastructure and academic support systems as well as institutional benefits in the form of recognition of the value of research and financial benefits in investing in research and incentives to researchers. The work illustrated strong evidence of North-South and South-South collaboration. There was a direct positive contribution to the scientific knowledge base, regional harmonisation, networking and indigenous doctoral programmes. But there appears to be little formal evidence for the development benefits of the three kinds of HE investment mentioned. Further, there are potential negative development effects of HE interventions, including ‘brain drain’ and inequitable, Northern-driven research partnerships. Finally, the internationalisation of universities, developments in ICTs, and private HE provision, particularly may have changed the landscape for HE interventions for development.

HE investments still lack a convincing theory of change that will lead to development outcomes and positive societal impacts in different contexts and sectors. The relationship between individuals, institutions and society is poorly understood with much of the emphasis being on outputs and little comparison between different approaches. For all these limitations, there are new and innovative tools being developed and tested for evaluating HE interventions, and considerable potential to adapt and apply a range of new methods and complete analysis of different kinds of interventions from other development areas. These can be built into new interventions from their outset, or used to analyse historical data sets provided by long-running programmes.

In conclusion there are lessons learnt from the literature review:

- The importance of process;
- The need to focus on learning;
- The benefits of using several methods;
- The need to include stakeholders in inclusive participatory ways, for design, effectiveness, implementation, impact and the coherence and relevance of the programme (duration, location, stakeholder involvement etc);

- The importance of initiating processes of reflection about assumptions and power; and learning by individuals, their behavioural changes and the contributions to institutional process and thinking including fellow mobility;
- Being clear about the definitions of impact, hard vs. Soft skills development objectives, on careers, publications, promotions, behaviours, adoption of outputs, perceptions, citations, grants, fundraising, spin-off;
- The need for ARD programmes to understand that different kinds of technologies have different social outcomes and that a technological solution may not be the desired solution every time;
- The importance of stakeholder participation and deliberation at several junctures in time and space to build an understanding of complexity in rural social dynamics. Lastly, stakeholder deliberations must reflect on power relations and expose them with an aim to redress them towards real and equitable partnership.

All capacity-development interventions are based on theories that are implicit in the minds of those who design and implement the interventions, because there is seldom a consensus among key stakeholders on the programme theory. ALINE's literature review has highlighted alternative and complementary participatory methods to incorporate notions about complexity and the multi-stakeholder nature of research, and its impacts on social change.

Key words: Capacity Development, Evaluation, Higher Education, International Development, Monitoring, practical lessons for application

Résumé

Le renforcement des capacités est important pour permettre à l'Afrique sub-saharienne (ASS) de faire concurrence dans les développements économiques et technologiques croissants au niveau mondial. Dans le secteur agricole, l'Afrique doit être en mesure d'atteindre la sécurité alimentaire et nutritionnelle et développer des systèmes d'innovation robustes pour les produits appropriés à sa propre utilisation. La nécessité de renforcer les capacités a mené à de nombreux programmes qui emploient une variété de modèles d'intervention. Beaucoup ont semblé être à la fois populaires et réussis. Toutefois, l'évaluation à plus long terme, s'est avérée difficile, avec les bailleurs de fonds étant particulièrement intéressés par le bénéfice visible sur leurs investissements. Ainsi, le regain d'intérêt dans le renforcement

des capacités pour le développement international a apporté avec lui un débat sur la meilleure façon de suivre de près et d'évaluer le renforcement des capacités, d'une manière efficiente et efficace. Il est maintenant reconnu que, même si l'individu est le centre direct de plusieurs interventions, le renforcement des capacités durable et à long terme est un processus complexe qui dépend d'une multitude de différentes parties prenantes et d'acteurs influents, nécessitant des modifications dans les systèmes et les changements dans les relations de pouvoir. Ce qui n'est pas clair est comment le perfectionnement des capacités contribue au développement international. Alors que les évaluations d'impact classiques saisissent les résultats les plus immédiats et pratiques de l'investissement dans l'enseignement supérieur, il y a aussi des avantages sociaux pour des interventions de l'enseignement supérieur qui ont trait à la sensibilisation social et au changement social, et agissent sur des périodes de temps plus longues. Cette étude examine la littérature et l'évidence pour comprendre si et comment les interventions de l'enseignement supérieur ES pour le développement sont en cours d'évaluation. Le travail reflète également l'évidence de l'impact. Cette étude met en évidence le taux d'achèvement réussi de la plupart des programmes de bourses dus en partie aux méthodes de sélection rigoureuses et équitables. Ce sont les programmes les plus courants qui doivent être évalués. Il est évident que les anciens boursiers appliquent ce qu'ils ont appris, participent dans la communauté et plus largement à travers leurs institutions et par la formation des autres.

Les consortiums doivent jouer un rôle réussi dans la mobilisation de nouvelles connaissances et des ressources et dans le transfert des connaissances concernant la recherche à d'autres (en particulier dans les partenariats des pays en développement). Il existe de plus en plus l'évidence d'un passage vers les produits locaux, les cours locaux, les programmes de doctorat, le leadership et les compétences, et la création des centres d'excellence. De nombreux membres ont mis en pratique les connaissances acquises et participent au niveau communautaire et plus largement encore. Il existe des preuves de renforcement des capacités avec succès des systèmes de soutien académique, des politiques et des infrastructures ainsi que des avantages institutionnels sous la forme de reconnaissance de la valeur de la recherche et des avantages financiers en investissant dans la recherche et les stimulations pour les chercheurs. Le travail a illustré des preuves solides de collaboration Nord-Sud et Sud-

Sud. Il y a eu une contribution positive directe au fondement des connaissances scientifiques, à l'harmonisation régionale, au travail en réseaux et aux programmes doctoraux indigènes. Mais, il semble y avoir peu d'évidence formelle pour les bénéfices du développement de trois types mentionnés d'investissement dans l'enseignement supérieur. En outre, il existe des potentiels effets négatifs du développement des interventions de l'enseignement supérieur, y compris la «fuite des cerveaux» et les partenariats de recherche inéquitables, dirigés par l'occident. Enfin, l'internationalisation des universités, l'évolution des TIC et la prestation de l'enseignement supérieur privé peuvent, en particulier, avoir changé le paysage des interventions de l'enseignement supérieur au développement. Les investissements de l'enseignement supérieur manquent encore une théorie convaincante de changement qui mènera aux résultats de développement et aux impacts positifs sociétaux dans différents contextes et secteurs. La relation entre les individus, les institutions et la société est mal comprise avec un grand accent mis sur les résultats et peu de comparaison mise entre les différentes approches.

Pour toutes ces limitations, il y a des outils nouveaux et novateurs mis au point et testés pour évaluer les interventions de l'ES, et le potentiel considérable pour adapter et appliquer toute une gamme de nouvelles méthodes et l'analyse complète de différents types d'interventions des autres domaines de développement. Ceux-ci peuvent être intégrés dans les nouvelles interventions dès le début ou utilisés pour analyser les ensembles de données historiques fournies par les programmes de longue durée.

En conclusion, il y a des leçons tirées de la revue de littérature:

- L'importance du processus;
- La nécessité de mettre l'accent sur l'apprentissage;
- Les avantages de l'utilisation de plusieurs méthodes;
- La nécessité d'inclure les parties prenantes dans une démarche participative inclusive pour la conception, l'efficacité, la mise en œuvre, l'impact et, la cohérence et la pertinence du programme (durée, lieu, implication des parties prenantes, etc);
- L'importance d'initier des processus de réflexion sur les hypothèses et le pouvoir, et l'apprentissage par les individus, leurs changements de comportement et les contributions au

- processus institutionnel et de la pensée, y compris la mobilité des individus;
- Être clair sur les définitions de l'impact, des objectifs de développement des compétences solides ou faibles ; sur les carrières, les publications, les promotions, les comportements, l'adoption des résultats, les perceptions, les citations, les subventions, les collectes de fonds, les avantages inattendus;
 - La nécessité de programmes d'ARD à comprendre que différents types de technologies ont différents effets sociaux et qu'une solution technologique ne peut pas être la solution recherchée à chaque fois;
 - L'importance de la participation des parties prenantes et de la délibération à plusieurs moments dans le temps et l'espace pour construire une compréhension de la complexité dans la dynamique sociale des milieux ruraux. Enfin, les délibérations des parties prenantes doivent refléter les relations de pouvoir et les exposer dans le but d'y remédier vers un partenariat réel et équitable.

Toutes les interventions de renforcement des capacités sont basées sur des théories qui sont implicites dans les esprits de ceux qui conçoivent et appliquent les interventions, car il y a rarement un consensus entre les principales parties prenantes sur la théorie du programme. La revue de littérature d'ALINe (Agricultural Learning and Impacts Network) a mis en évidence des méthodes participatives alternatives et complémentaires pour intégrer des notions sur la complexité et la nature de la multi-partie prenante de la recherche, et ses impacts sur le changement social.

Mots clés: Enseignement supérieur, développement des capacités, développement international, suivi, évaluation, leçons pratiques d'application.

Background

Capacity strengthening¹ has an increasingly prominent position in international development, enabling Sub-Saharan Africa to compete with growing global technological and economic developments. In the agriculture sector, Africa must reach a position to achieve food and nutritional security by overcoming the technology gap, through strengthening regional and national innovation systems to develop and adapt products appropriate for its use.²

¹ Also known as capacity building or capacity development

^{2&4} Conway, G. et al. "The Montpellier Panel Report Africa and Europe: Partnerships for Agricultural Development" 2010

Furthermore, Africa is expected to add 122 million people to its labour force between 2010 and 2020. By 2035 the continents workforce will exceed that of any nation including China and India. Africa has the potential to create 54 to 72 million more stable wage-paying jobs with much of that growth coming from manufacturing, agriculture, and retail & hospitality³. The dependency ratio – *the number of young and retired people that each working person has to support* – is set to decline and be on a par with North America and Europe by 2035⁴.

Agriculture is set to create 8 million stable jobs by 2020 (at current trends) but could add a further 6 million if the continent accelerates the development of this sector⁵. This could come from expanding large-scale commercial farming on uncultivated land and shifting from low-value grain production to more labour-intensive and higher-value-added horticultural and biofuel crops.

Despite a noticeable withdrawal of donor investment in many aspects of tertiary education across Africa over the past two decades⁶, the need to strengthen capacity has provided an opportunity for numerous programmes employing a variety of models of intervention. Many have appeared to be both popular and successful. However, evaluation over the longer-term has proved difficult, with donors being particularly interested in the visible return on their investments. **So the renewed interest in capacity strengthening for international development has brought with it a debate about how best to monitor and evaluate capacity strengthening, efficiently and effectively.**

There is increasing awareness that even though the individual is the direct focus of many interventions, long-term and sustainable capacity strengthening is a complex process dependent on a range of different stakeholders and influencing players, requiring changes in systems and shifts in power relationships. What is not clear is what and how capacity development contributes to the wider community in the

³ Africa at work: Job creation and inclusive growth McKinsey Global Institute http://www.mckinsey.com/insights/mgi/research/africa_europe_middle_east/africa_at_work

⁴ The world at work: Jobs, pay and skills for 3.5 billion people, McKinsey Global Institute, June 2012 http://www.mckinsey.com/insights/mgi/research/labor_markets/the_world_at_work

⁵ Lions on the move: The progress and potential of African economies, McKinsey Global Institute, June 2010 http://www.mckinsey.com/insights/mgi/research/productivity_competitiveness_and_growth/lions_on_the_move

international development context. While conventional impact evaluations capture the more immediate and practical outcomes of Higher Education investment, there are also societal benefits to Higher Education interventions which relate to social awareness and change, and act over longer time periods.⁷

Through examining the contributions of these initiatives to change, successful practices can be identified and fed into a broader understanding of capacity development.

The purpose of this paper is to share our understanding of challenges and opportunities around measuring the impact of Higher Education (HE) interventions, to illustrate some of the methods that are being pursued by others in the field as well as to use examples from our own work to identify opportunities to capture the value of HE interventions for wider society (including governments and donors), institutions and individuals.

This paper summarises four main bodies of recent experience:

1. A Conference held on 19-20 March 2012 in London between the London International Development Centre (LIDC), the Association of Commonwealth Universities (ACU) and the Institute of Education (IOE), which established dialogue between the HE sector, the international development sector and the evaluation community⁸
- . 2. A systematic literature review commissioned by ALINe to review approaches in evaluating HE capacity development interventions for agricultural development.⁹
3. An ex post evaluation commissioned by ALINe on the Rothamsted Research African Fellowship Programme.¹⁰

⁷ LIDC Report “Examining development evaluation in higher education interventions: a preliminary study” By Charlotte Creed with Hilary Perraton & Jeff Waage.

⁸ Conference Materials: http://www.lidc.org.uk/news_detail.php?news_id=149.

Short Conference Summary and Recommendations: http://www.lidc.org.uk/_assets/Conference%20Summary%20and%20Recommendations%20FINAL.pdf

The full summary: http://www.lidc.org.uk/_assets/Conference%20report%20Impact%20of%20HE.pdf

The baseline study on evaluating HE interventions that LIDC developed for the conference can be accessed from: http://www.lidc.org.uk/_assets/LIDC%20Higher%20Education%20study%20final.pdf

⁹ Lorenzo, S.R. (2011) ‘Evaluating Individual Approaches to Capacity Development: A Literature Review’, ALINe Working Paper (forthcoming), <http://www.aline.org.uk/resources/alineworkingpapers>

¹⁰ Mann Chaussalet, J. (2012) ‘Evaluation of capacity enhancement fellowships for agricultural development: A case study of the Rothamsted International African Fellows Programme’, ALINe Working Paper (forthcoming). <http://www.aline.org.uk/resources/alineworkingpapers>

4. ALINE's relationship and experience with RUFORUM working on establishing an effective and practical monitoring and evaluation system for its regional programmes.

There is a long history of investment in HE as a contribution to international development. These include university training through scholarship programmes and capacity building projects with universities. Capacity building elements have also been incorporated into research projects and within research consortia. Some but not many have been evaluated systematically.

Donors, northern and southern research organisations have different degrees of voice and decision-making powers in the design, implementation and evaluation of capacity development programmes. Similarly, those affected by the research outputs – e.g. different kinds of small-scale farmers in the case of agriculture research and development (ARD) – and other stakeholders need to participate in these processes.

Recent literature in capacity development has highlighted the need to incorporate dialogue into the research cycles to avoid schemes being dominated by donor agendas. Openness, devolution and dialogue by donors must be matched by recipient organisations being clear in their purposes and agenda, and having the courage for ‘hard talk’ when necessary. The importance of building networks and relationships in research in order to create ‘knowledge communities’ cannot be underestimated for sustaining the benefits and ensuring that there is adequate time for those benefits to begin to improve the skills, opportunities and successes of individuals and institutions.

It is hoped that better evidence of the value of HE interventions can lead to more effective policies and interventions in support for HE for development.

Higher Education Interventions in Development

What do we mean by capacity development? Often the terms ‘capacity building’, ‘strengthening capacity’ and ‘developing capacity’ are used with different meanings. In the late 1980s and early 1990s capacity development emerged recognising that the failure of previous approaches (“technical assistance”, “technical cooperation”) ignored existing capacities. Donors and organisations began to understand that capacity development is an internal process that involves the main actors

taking primary responsibility for change processes^{11,12}. Yet this internal emphasis does not mean that external support is not useful or desirable, but that external support should catalyse and facilitate internal processes of change.

A useful definition for capacity is that used by the OECD: “*the capability of people, organisations and society as a whole to manage their affairs successfully*” and capacity development is, “*processes whereby people, organisations and society as whole, unleash, strengthen, create, adapt and maintain capacity over time*”¹³. This definition illustrates that capacity development occurs in at least three levels: the individual, the organisation and broader society. Different development organisations place varying emphasis on different levels. There are many dimensions of capacity development and systemic notions of capacity development force us not to restrict our analysis to only one dimension in isolation. A systemic approach requires us to understand the strong linkages that exist between the fellow, his or her organisation and broader society.

In the literature on HE interventions and the recent LIDC conference, there are three broad categorisations although there is often overlap between these three broad categories;

1. Education and training initiatives. Courses designed to improve capacity are run in developing and industrialised countries have been in existence for many decades now. High income country governments have also provided professional training for students from developing countries. Donors have also sought to raise capacity such as USAID in agriculture by running training courses and supporting students at American universities. Most of these activities fall into the following categories;

- a) Specialised training courses with a clearly defined purpose and target audience
- b) Open ended scholarship programmes run with broad objectives and open to applicants in a range of disciplines

¹¹ UNDP (2009) Capacity Development: a UNDP primer, New York: UNDP

¹² Simister, N. and Smith, R. (2010) ‘Monitoring and Evaluating Capacity Building: Is it really that difficult?’ INTRAC Praxis Paper 23.

¹³ OECD (2006) ‘The Challenge of Capacity Development: Working towards good practice’, p12, Paris: Organisation for Economic Development and Cooperation.

- c) Focused scholarship programmes restricted to particular institutions or disciplines
- d) Distance-learning scholarships introduced as an alternative to scholarships that require scholarship holders to travel in order to study.

2. Consortia and networks. Universities have been supported to link departments that have shared interests in support of research and teaching but also of building capacity in those partner institutions. This is apparent at regional levels (e.g. RUFORUM) or internationally such as Britain's HELPIS programme (1981-2006) and more recently DELPHE which promotes partnerships between universities and other higher education institutions HEIs working on collaborative activity linked to the MDGs in priority countries of the UK's Department for International Development (DFID). There are others such as the Foundation Partnership in Higher Education, past Carnegie support and even more recent initiatives such as the Wellcome Trust's African Institutions Initiative¹⁴. Tertiary Education for Africa Mechanism (TEAM Africa)¹⁵ and the African Network for Agriculture, Agro forestry and Natural Resources Education (ANAFE)¹⁶ are also similar consortia based initiatives.

Donor funded consortia on development issues such as health, education, agriculture and other sectors are likely to have capacity development as a subsidiary objective to their main purpose. This is true where there are consortia which include advances international or national research institutions in the region or further afield also.

3. Supporting improved institutional capacity. Projects or programmes have been run with the direct aim of improving and increasing capacity within institutions individually or collectively and was the rationale for increased investment in universities in Africa in the 50's and 60's, with collaborations with universities in the US and the UK, exchange of staff, scholarships for staff and potential senior staff. Recent activity includes strengthening networks of universities, more narrowly

¹⁴ <http://www.africaninstitutionsinitiative.org/>

¹⁵ TEAM-Africa - Association of Public and Land-grant Universities www.aplu.org/document.doc?id=3468

¹⁶ <http://anafe-africa.org/?q=content/about-anafe>

Pinto, Y.

focused on ICTs or more broadly focused on regional level initiatives such as harmonised student loan policies.

The value of universities is clearly beyond their contributions to economic growth, they create open spaces for challenging and understanding social change over periods of time, they include effects on the intellectual outlook of individuals and organisations especially since these changes may be manifested several years in the future.

However different interventions will have different conditions for success and are likely to require different types of evaluation and methods of measurement.

There is also a recognition that the decline in investment in HE relative to basic education has been associated with the view that HE interventions are comparatively poor value for money, inequitable and not development focused.

How can we change this situation, what can we do and who needs to take the initiative including what do we hope to achieve, for whom, over what timeframe?

Initial approaches saw these three categories as separate entities, yet these three levels influence each other in a fluid way. Access to resources and experiences that can develop individual capacity is largely shaped by organisational and environmental factors, which are in turn influenced by the degree of capacity development within each individual¹⁷. Our approach must therefore be systemic¹⁷. Systemic approaches permit us to explore the relationships between learning and change across the three levels (individual, organisational, societal) in an integrated way¹⁸. This systems approach does not only refer to the fact that we need to consider the relational nature of individuals, organisations and society, but also that social change is complex, and that we cannot keep the simplistically linear assumptions of technical cooperation¹⁹.

¹⁷ Clarke, P. and K. Oswald (2010) ‘Introduction: Why Reflect Collectively on Capacities for Change?’ IDS Bulletin 41(3): 1-12.

¹⁸ Taylor, P. and Clarke, P. (2007) Capacity for a change. Document based on outcomes of the ‘Capacity Collective’ workshop. Capacity Collective Workshop, Dunford House.

¹⁹ Akum, R. J. (2007) Knowledge Production, Sharing and Use: Programming Capacity Valorisation Appendix C, Scoping Paper 1, p3, Capacity Collective Workshop, CODESRIA, Dunford House.

Towards a Systematic Understanding of Capacity Development

In Morgan's words, capacity is an 'emergent property or an interaction effect. It comes out of the dynamics involving a complex combination of attitudes, resources, strategies and skills, both tangible and intangible. It emerges from the positioning of a system within a particular context.²⁰

What does this systems approach entail in the analysis of individual approaches to capacity development?

Individual approaches to capacity development focus on individual learning as a form of capacity development. This individual acquisition of skills, experience and knowledge can occur formally, through education and training, or informally, for example through doing and observing. These individual approaches assume that:

- Individuals possess knowledge, skills and attitudes which reflect their experience and training.
- When individuals share their knowledge, skills and attitude with colleagues and these become embedded in group norms and processes, it can be said that they have become part of the groups' capacity.
- And when the individual and group capacities become widely shared among the organisations' members and incorporated into management systems and culture, they become organisational capacities²¹.

However, do individual approaches to capacity development need to take a systemic approach? Do schemes like fellowships acknowledge and integrate into their programmes the relational nature of individuals, organisations and society or do they leave organisations and society out of the equation?

Individual approaches to capacity development encompass developing the skills of individuals and, in doing so, contributing to change in organisational procedures. Individuals may take their acquired skills when they leave the organisation, yet if the scheme was successful, new systems and procedures have become integral to the operations of the organisations. Thus individual approaches can only be successful if new approaches

²⁰ Morgan, P. (2006) *The Concept of Capacity*, p7, Brussels: ECDPM.

²¹ ECPDM (2003) 'Learning about capacity development through its evaluation: experiences from Africa, Asia, Europe and the Americas', *Capacity.org* 17, p4.

(acquired or developed by individuals) are institutionalised²² (Kristjanson, Lilja *et al.*, 2008: 10). Despite focusing on the individual learning, these approaches are more likely to be successful if they are framed within a systems approach.

Unfortunately, some individual training programmes that ignored the organisational-level capacities of partner organisations can have adverse effects; although selected individuals enhanced their skills, some projects placed harsh ‘demands on already weak and overstretched organisations’²³. DFID’s evaluation of the Renewable Natural Resources Research Strategy showed that putting all resources into training researchers to certain levels improves research outputs and networks, yet does not contribute to improving project and resource management systems and the definition of new research strategies (*ibid*).

The organisational capacity of partners should be a priority as should understanding the circumstances in which individual and organisational interests converge or compete against each other. It is necessary for programmes to show an understanding of the importance of management infrastructure, technologies and strategic and policy capabilities at the organisational level in order to create a lasting impact in research systems.

When monitoring and evaluating individual approaches to capacity development, our analysis cannot therefore separate individual learning from organisational and institutional learning. For example, the Rockefeller Foundation places a great emphasis on institutions in its capacity development programmes, stating that ‘building skills systematically across local organisations and among organisations in different countries’, catalyses ‘an environment of inquiry, entrepreneurship and experimentation’ that in turn, ‘makes individuals and organisations more effective’²⁴. In fact, organisational learning theory acknowledges that organisations can learn independently of the individuals within them. Changes in rules, procedures, culture or structures within an organisation can lead to organisational learning²⁵.

²² Kristjanson, P., Lilja, N., et al. (2008) ‘Rethinking Impact: Understanding the complexity of poverty and change. Key issues discussed at the workshop.’ ILAC Working Paper 6: 1-16.

²³ DFID (2006). ‘An Integrated Approach to Capacity Development, Renewable Natural Resources Research Strategy’, p3, London: Department for International Development (DFID).

²⁴ Lewinger Moock, J. (2004) ‘Rockefeller Foundation: How We Invest in Capacity Building’, p3, Rockefeller Foundation.

²⁵ Taylor, P. and Oswald K. (2010) ‘A learning approach to monitoring and evaluation’, IDS Bulletin 41(6): 115.

In short, when conducting a capacity needs assessment for a fellowship, or when designing or evaluating a capacity development programme with a strong individual learning component, we must take into account the importance of exploring the relationships between learning and change across the three levels (individual, organisational, societal) in an integrated way. The programme must test its assumptions on how individual change will lead to positive changes in the organisations concerned and ultimately in society. Similarly, a programme design must show understanding of how social dynamics within organisations shape, constrain and facilitate individual learning and capitalise on this. Therefore, evaluations of fellowships must include assessments of whether organisational learning has been addressed, regardless of whether its efforts remain focused on the individual.

Do schemes like fellowships acknowledge and integrate into their programmes the relational nature of individuals, organisations and society or do they leave organisations and society out of the equation?

Or perhaps:

Did the individual approaches (acquired or developed knowledge by individuals) become institutionalised in some way?

Hard and Soft Capacities

There are two dominant models in capacity development schemes, whether they support the development of ‘hard’ capacities (infrastructure, technology, finance and so on), or ‘soft capacities’ (capacity to manage knowledge, develop organisational procedures and so on).

Similarly in individual approaches to capacity development, there are two tendencies in professional development: one that focuses more on ‘hard science’, for instance, learning new technologies or procedures (for example, the Borlaug programme or the AFP), and another that puts greater emphasis on ‘soft science’ skills, including networking capacity, leadership, confidence and resilience (a good example is the AWARD fellowships). It is extremely important to tailor capacity development schemes to the specific context and social dynamics of the organisations involved.

The most extreme differences in approach are seen in areas of ‘high technological research’, including science focused

fellowships such as the Borlaug Fellowships, Marie Curie Fellowships, OECD Fellowships, and Wellcome Fellowships where the aims emphasize the learning or use of technologies, and advancement of scientific research. Evaluation indicators for such programmes include “research outputs”, “research achievements” and “publications”. However, other fellowships with an influence from social sciences have a much stronger focus on personal development, where evaluation indicators include aspects of confidence, esteem, leadership and management (CGIAR AWARD Programme), and social justice (Ford Foundation). Others focus at assessing impact on the three levels given above with more or less success (International Foundation for Science, UK Commonwealth Scholarships).

While scientists tend to focus by definition on “hard” capacity development, there is a considerable amount of evidence suggesting that investment in ‘hard’ capacities rarely leads to improvements in organisational performance unless soft capacities such as strategic leadership, programme and process management and networks are nurtured as well¹⁹. Individual approaches to capacity development such as science research fellowships need to develop not only ‘hard’ science skills related to scientific research, teaching the use of new technologies and methods, but also develop ‘soft’ science skills that are equally important: the capacity to network and build relationships with others, the strengthening of leadership skills, the fostering of confidence and resilience²⁶.

The difference between hard and soft capacity components is crucial in individual approaches to capacity development. Not only does one need to focus on the organisational capacities – its resources and management- but on the hard and soft skills acquisition by the individual and the implications for his/her organisation and environment. Thus the success of individual approaches to capacity development depends on the ability that programmes have in providing training or research exchange opportunities that are part of a multi-level approach to organisational capacity development. They should be tailored for the specific context and social dynamics, and must build-in incentives to apply the new skills, articulate personnel development, and link this with team performance and overall organisational efficiency.

²⁶ Ofir, Z., Van Wyk, B., et al. (2008), ‘Comparative Evaluation of the G&D-Rockefeller and Borlaug Women in Science Fellowship Programs’, Working Paper 49.

Another aspect to be considered is the extent to which capacity development is influenced by hierarchical or “**power**” **interactions**. These need to be taken into account where possible and can occur at all three levels: individual, organisational, and wider society. For example, those who hold the funds, or the organisations who are gatekeepers of these funds, wield power that might shape the partnerships being developed and this can lead to long-term processes that deliver intangible results which focus solely on accountability to donors rather than achieving developmental change. Such influences can be manifest in terms of unequal access to information and opportunities to publish or author articles. Alternatively, northern partners may have a disproportionate influence on decision-making in terms of the research agenda, and project negotiation in general.

In parallel to the promotion of **equitable partnerships and consortia**, the promotion of networks and relationships is crucial in creating capacity. They depend on the construction and expansion of personal contacts, and rely highly on trust and informality. Despite the mounting evidence that networks are a fundamental channel to exchange knowledge and ideas, development programmes often do not assign them formal recognition.

However, the review of literature has shown that this is not the case for Individual approaches to capacity development such as fellowships. Capacity development in Agricultural Research and Development has understood for some time the **power of professional (and personal) networks** and many ARD fellowships include specific activities aimed to strengthen alliances between professionals such as exchanges, conferences, workshops, and informal events.

Impact Measurement and Evaluation Experiences Todate

In March 2012, the London International Development Centre (LIDC) and Association of Commonwealth Universities (ACU) organised a study and a conference to explore past and current efforts to measure the impact of HE interventions for development, what they tell us and where they can be improved²⁷. This is part of a larger initiative to understand how HE for development targets should be built into post-2015 goals setting, and the kinds of indicators that will be needed for this.

²⁷The baseline study on evaluating HE interventions that LIDC developed for the conference can be accessed from: <http://www.lidc.org.uk/assets/LIDC%20Higher%20Education%20study%20final.pdf>

Pinto, Y.

This initiative was led by the Institute of Education, a member College of LIDC.

The initial study²⁵ illustrated that only a fraction of relevant programmes have clearly articulated development objectives and procedures for determining whether these objectives have been achieved: only 17 of 67 programmes investigated met the criteria for the study. The small sample made it difficult to draw broad conclusion about the effectiveness of HE interventions for development.

Most of the evaluations, which are all donor-sponsored and donor-defined, try to assess whether projects have been successful, in their overall development approach and in promoting certain widely-held donor principles. These include becoming more demand-driven, successful in promoting a sense of ownership in the south or reducing gender discrepancies. These principles are incorporated into interventions to differing degrees and in differing ways, e.g. self-initiated proposals, equitable selection procedures which positively target women and people from less conventional academic backgrounds or disadvantaged groups, and sharing of costs and responsibilities. Evaluators are assessing these against explicitly stated objectives in the original or later reshaping of the programme design or else looked at retrospectively in the evaluation. Most of the evaluations focus heavily on quantifying a programme's outputs, such as the number of research projects/training sessions/PhD or Masters in specific subject areas, the geographical reach of the project and measuring institutional links within a region.

Donor preoccupation with these ownership principles has also revealed unexpected outcomes. Evaluation of the IUCEA-DAAD programme²⁸, for example, found the concept of stakeholder consultation met with some resistance among its East African partners. In others, programme targets for gender balance and more representation from disadvantaged groups remain doggedly disappointing or else the equitable selection procedures of the programme are in stark contrast to realities

²⁸https://docs.google.com/viewer?a=v&q=cache:pV7_ockVNy0J:eu.daad.de/imperia/md/content/eu/downloads/drittland-kooperationen/veranstaltungen/hansert_the_iucea_-_daad_partnership_for_reforming_quality_assurance_in_east_african_he.pdf+IUCEA+DAAD&hl=en&gl=uk&pid=bl&srcid=ADGEESiN6jAA7X7R6M qYSz3e2K0qNVn TArZ09A22yuOZy7 T6KrE0N0T9c7lfMXVcVUt QVfm4IjUHg kZWoNkzM46JHU9r9ySfLNHWIucta Y9s32Pl9 THv9ehBbJSahn oi7Uujte tu&sig=AHIEtBQi2QfGFwOTRdrGnpYiqlCcGqBVGw

in a more restrictive society. The SIDA-SAREC²⁹ evaluation found that although the shift to self-managed research funds had positive outcomes, it nevertheless raised difficult questions about efficiency and accountability. DFID allows for its research consortia to be led by or consist entirely of developing-country partners. In practice, however, evaluations have found that expertise in proposal writing and budget and administrative support, key to successful bidding, were not evident in most African partners compared to the UK ones. The principle of local level ownership and project definition, then, is there in theory but not always in practice. It is a process which needs continuous capacity building with additional bridging activities.

A recurrent issue in the evaluations is an imbalance between beneficiaries and funding agencies in terms of interests (e.g. funding agencies, beneficiaries, North vs. South), objectives and/or perceptions of success.

Some evaluations stress the need for a balance between achieving success in the beneficiaries' terms and generating outputs that are useful to donors, and also a balance between a view of development that values development as process, rather than the achievement of pre-specified targets and goals. Just as we may need to ask about the goals of projects or programmes, so we need to ask about the values of those doing the evaluation and ask whether new insights will be gained by different types of evaluators and evaluations.

The drive to demonstrate results is leading funding agencies to ask for more evaluation about the impact of their interventions. This has raised certain methodological difficulties:

1. The desire to measure impact does not yet appear to be matched by a range of methodological tools in mainstream use.
2. Some programmes are being evaluated in terms not built into the original objectives or where objectives about 'capacity building in HEIs' are so generally stated that evaluators are retrospectively bolting on an assessment of impact as just one component in a wider-ranging evaluation where, to give it its due, impact should be the sole focus of an evaluation.

²⁹<http://knowledge.cta.int/en/Dossiers/S-T-Policy/EU-agricultural-S-T-dialogue/S-T-Organisations-Web-resources/S-T-in-Sweden/SIDA-Department-for-Research-Cooperation-SAREC>

3. The difficulties of assessing impact and assigning attribution in multidimensional programmes or which have subsequently become subsumed under other activities and it therefore becomes difficult to identify an intervention's separate and distinct contribution.
4. Prematurely measuring impact in relatively recent interventions.

Trends in globalisation and the internationalisation of education are often taken into account in HE interventions and this is reflected in some evaluations. For example:

- a) Measuring the internet profile and publications of academic beneficiaries/HEIs/departments by means of simple indicators (tracing citations & sites of citation to measure the degree of influence);
- b) Assessing the effectiveness of credit transfer systems in promoting student mobility;
- c) examining the state of progress in the establishment and effectiveness of regional networks and programmes or systems, e.g. the standardisation of Quality Assurance system within a region;
- d) Number of student/academic/research and support staff exchanges;
- e) Volume of publications in international journals or participation in/management of international conferences;
- f) Evidence about the development and usage of HEI/regional research area websites and pan-regional/African publications;
- g) Measuring whether access to international journals, associations, special interest groups is prompting new research or practices.

Where once the preoccupation in scholarships schemes was to ensure a good rate of return and avoid brain drain, we are now seeing some scholarship programmes, e.g. AusAID, also embracing the idea of increasing skills for migration through projects with the explicit aim of preparing participants for the global nursing and teaching markets.

The evaluations studied reflect broader trends in the history of technical assistance for capacity building in higher education. For example, the historical broadening out of capacity building projects at individual level, such as scholarship schemes, towards programmes for institutional development and research consortia

Challenges in Evaluating the Impact of ARD Programmes and their Relevance

General Characteristics of the Evaluations Reviewed

has inevitably impacted on the scope of evaluation and what and how much is being evaluated, how it is evaluated and who is involved in evaluation. The distinction between the types of interventions has in some cases become blurred in some multidimensional programmes (which combine aspects of some or all of them) and the evaluations, in turn, are attempting to measure these more complex interventions.

There are important challenges that exist when assessing the indirect impact of fellowship schemes' activities on broader society. For example, many agricultural research fellowships ultimately aim to have an effect on agricultural productivity and small-scale farmers' livelihoods. However measuring impact is no easy task as ARD systems are complex and linear assumptions about causality and change are not applicable. **Appendix I captures some of the challenges and some of the methods that are being pursued to overcome them.**

The evaluations examined were mostly led by developed country evaluators (internal or external) although there is some, limited, evidence that the findings are acknowledged by those involved in conducting the evaluation itself: ***mixed-team field visits*** (e.g. IUCEA-DAAD where one European external HE expert is accompanied by two East African experts) and ***user-end case studies*** (African teacher educators rather than European evaluators assessing their experiences of the adoption and practice of TESSA Open Educational Resources³⁰(OERs) in their institutions); ***respondent validation*** (where key southern partners are asked to comment critically on the draft of an evaluation as a pre-requisite before its completion, e.g. SIDA-SAREC). Data collection in all the evaluations draws on the ubiquitous ***stakeholder consultation***. This varies from the generally stated and, apparently token, short 'lists of those consulted' to more comprehensive and systematic consultation, e.g. the development and careful piloting of questionnaires for a wide range of different stakeholders, successive Delphi-method group interviewing and the involvement of a wider range of complementary stakeholders external to the immediate HEI project.

Of the evaluations studied here, ***the general picture is a focus on the outputs of a programme or project, far more than effects and impact.*** It is not too difficult to identify and get evidence on outputs despite the fact that many of the donors

³⁰ <http://www.tessafrica.net/>

Pinto, Y.

‘Terms of Reference’ (TORs) also want assessments of effects and impact, which are much further down the road, and with far more intervening variables that get in the way of ascribing an outcome to a particular earlier output.

As previously mentioned, there is a distinct sense that an assessment of impact has been bolted on to the evaluation as just one aspect of a larger evaluation and without, seemingly, the methodological tools to do so other than evaluator(s) consider impressions about impact. The definition of impact also seems to vary. If we were to take scholarships as an example, then one might classify university completion rates as an output measure, the amount of promotion at work within five years as an effect measure and the role of scholars and their effect on society over 20 years as an impact measure. Some evaluation, on the other hand, may consider these effects to be programme impacts.

In identifying evaluations for the LIDC study, scholarship evaluations were more readily available than other types. Included in these are two, the Ford and Commonwealth Scholarships³¹ Commission, which are assessing impact on the basis of, respectively, a 40- and a 50-year programme which is very rare. The new impact studies of HE interventions, which do include specific and rigorous impact methodologies, are in the pipeline but not yet publicly available or completed.

Methodology and Sources of Data

The small sample of evaluations here include; *expert reviews*, *peer reviews*, *internal progress reports* and *case studies*. Most of them are expert reviews, which is in contrast to traditional peer reviews where, for example, a scholar would focus on the quality of, say, the research outputs of a research capacity building intervention. Expert reviews typically have broader terms of reference, looking at a programme’s management and also assessing evidence of outputs, outcomes and impact, and often at programme, project and field levels. They tend to be both backward-looking (assessing progress, impact and relevance of past actions) and forward-looking, strategic in perspective in order to understand future challenges and inform next stages in programme or wider strategic planning. Typically, these expert reviews rely on the expertise of the evaluator(s) and their use and synthesis of a combination of a wide range of data sources to inform the evaluation. The latter

³¹ <http://www.intscholarships.com/region/africa>

tend to include some or all of the following: **document analysis** (earlier evaluations and reviews, annual reports, implementing guidelines, letters of agreement, annual reports at project and programme levels); the **use of monitoring data** at both project and programme level, including **cost data**; face-to-face, telephone **interviews** (of individuals and sometimes groups of direct beneficiaries, project and programme leaders and other stakeholders); **questionnaire surveys** of beneficiaries and other stakeholders, and **country visits**. Country visits can include field visits, case studies or the use of pre-existing case studies.

Evaluation of scholarships have used **alumni tracer studies** (postal and online surveys, sometimes backed up with in-depth case studies of individuals, visits and interviews) in order to look at evidence on effects and impact. The studies are of varying complexity ranging from the simple (What did you gain from the scholarship? Where are you now and what are you doing?) to more complex designs (involving detailed analyses of the career paths of different cohorts of alumni and examining the relationships between variables such as course studies, course levels, location of study, completion rates, rate of return, gender, age, types of jobs, changes in work remuneration and status, e.g. Ford's IFP³²). The robustness of conclusions from tracer studies is, reduced where only a small proportion of alumni respond. Most of the alumni studies come to the conclusion that a self-report survey alone is insufficient to measure effects and impact and there is an acknowledged additional need to drill down into specific groups, regions, sectors of employment to assess impact more thoroughly and add validity and reliability to self-reported impact findings.

There is also evidence of attempts to measure impact within the evaluations in other ways, e.g. small **case study impact** evaluation from an end-user perspective (e.g. TESSA, AGORA³³). In larger, multidimensional project evaluations, impact is explored as one area among others (e.g. effectiveness, efficiency) and at a rather impressionistic level (based on anecdotal accounts from project leaders) or the anticipated impact (based on the evaluators' retrospective assessment of the inputs and effects of the programme). In this sample of evaluations, there has been no use of **planned statistical measures** of impact evaluation (e.g. randomised control trial, propensity score matching, and regression discontinuity design).

³² <http://www.fordifp.net/>

³³<http://www.aginternetwork.org/en/>

Qualitative and Quantitative

Many of the evaluations examined use mixed qualitative and quantitative methods, although qualitative narrative accounts dominate. Where quantitative methods are used, they are usually simple number and percentage calculations in simple questionnaires, although some of the alumni tracer studies attempt greater depth by exploring a variety of variables. The USAID study of the ATLAS³⁴ (AFGRAD) scholarship programs is the only example of sample surveying on a large scale. There are no natural experiments or randomised trials. There appears to be a strong divide between qualitative or quantitative methods, deep divides between some quantitative and some qualitative researchers, a dearth of quantitative approaches.

Evaluative Frameworks

There are examples of evaluators working within donor-prescribed evaluative frameworks. The SIDA, ADB and VLIR-UOS³⁵ programmes, for example, draw on the Organisation for Economic and Cooperation Development/Development Assistance Committee (OECD/DAC) evaluation criteria of efficiency, effectiveness, quality, sustainability, development relevance, impact, programme design and management. In the VLIR-UOS programme, for example, each criterion was accompanied by a set of indicators. The impact criterion had the following set of indicators (although these might, depending on the definition of impact used, be seen as effects):

1. Active engagement of academic/administrative staff and students in planning/developing initiatives;
2. Evidence of changed approaches/attitudes to learning and teaching integrating ICT-based practices (e.g. engagement by academic staff in use of e-learning);
3. The library provides access for staff and students to relevant electronic and print information resources to support current university learning, teaching and research.
4. Change in levels of use of library resources and range of resources used (impact of effective library management system);
5. Spread of and changing levels of student skills and confidence in ICT use;
6. Specific research activities, dependent on access to e-information or data, enabled by ICT provision;

³⁴<http://www.atlascorps.org/>

³⁵http://www.scholarships.vliruos.be/index.php?language=EN&navid=479&direct_to=Overview%20Eligible%20Programmes

7. User focused website, meeting specific information needs.

Costs

There is little focus on costs beyond simple cost analysis (identifying and making sense of a project/programme's costs). Two exceptions include the Commonwealth Scholarship Commission³⁶ (CSC) and the ADB Japan Scholarship Programme³⁷ which both examine cost-effectiveness: CSC has looked at the comparative cost of producing a graduate through its conventional scholarships and through distance-learning scholarships. **ADB has compared the per graduate costs of scholarships delivered by different partner higher education institutes in different countries, e.g. \$19,000 in Singapore, \$3, 000 in Japan, \$45,000 in Australia.**

No cost-benefit study was found in the sample evaluations. One early cost-benefit analysis of scholarship programmes has been identified: In the 1970s an East African and American group collected and used salary data on alumni of local, eastern bloc, Western European and American scholarship programmes in order to undertake rate of return analysis (Maliyamkono *et al.*, 1982³⁸). Although there is extensive human-capital theory literature on the costed benefits of higher education generally, with this exception this has not usually disaggregated results in terms of local versus international study.

Reliability and Trustworthiness Issues

Few evaluations discussed explicitly their methodology in terms of reliability and trustworthiness. One recurrent acknowledged problem is the unreliability or inefficiency of monitoring systems at both the local project and central programme level; baseline data is not always available, complete or up-to-date and it is often conflicted at the different levels.

The heavy reliance on expert reviews raises certain credibility questions. The expert's credibility relies on two things: first, the perceived fairness and competence of the expert(s) and being demonstrably independent; second, the level of detail of the data sources available. The independence issue is complicated by the participation of evaluators with a close relationship with

³⁶<http://www.atlascorps.org/>

³⁷<http://www.adb.org/site/careers/japan-scholarship-program/main>

³⁸ Maliyamkono, T. L., Ishumi, A. G. M., Wells, S. J. and Migot-Adhola, S. E. (ed.) 1982 Training and productivity in eastern Africa. London: Heinemann

Pinto, Y.

the agency and also the small international community of development evaluators.

There is also the question of whether expert review is an adequate approach as projects become more multidimensional (as in research consortia and institutional capacity building) and as donors are asking for wide-ranging evaluative measurements (input, effects, impact, sustainability, development relevance, efficiency, quality, ownership, management, equity, backward and forward-looking evaluations).

The task of the evaluator(s) can range from assessing the quality of the research conducted to estimating its impact in developing countries and also include suggesting improvements in the management of the programme. Few can undertake all such tasks at an institutional level with the requisite methodological rigour. It is little wonder the evaluations are littered with rather plaintive or defensive remarks made by the evaluator(s).

The following is a summary of the methodologies (and some of their identified perceived strengths and weaknesses) as used within the three types of interventions:

Education and Training

Scholarship programmes naturally depend on tracer studies to assess their outputs, effects and impact. Without them, as in the case of the AUSAid expert review³⁹, evaluation can only operate at the level of anecdotal evidence about alumni promotions and influence in a small sample. Long-established schemes can undertake larger, sometimes rigorous random sample surveys and examine differences, such as female participation, between, say, earlier and later cohorts, or differences in quality or costs among different partner HEIs. Alumni surveys depend on the success in tracking or getting responses from alumni and responses rates can vary considerably. They are also heavily dependent on the quality of monitoring systems and, where the monitoring systems are split between programme- and project-level, the degree to which these split-site systems are compatible or consistent. This is clearly an area of weakness. Most evaluators acknowledge the need, either in their evaluation or in recommendations for the future, for additional in-depth case studies of particular

³⁹AUSAID, 2007 „Completing a quality at implementation report (QAI) , *Guidelines (unclassified)*, Canberra: AUSAID, accessed 1st February 2012 at http://www.spc.int/hiv_index2.php?option=com_docman&task=doc_view&gid=462&Itemid=148

scholars or departments (where, say, the scholarship aimed at group training of professionals in a particular workplace) and input from the HEI partners and employers.

Links and Consortia

The complexity and multidimensional nature of links and consortia capacity building programmes clearly present considerable evaluation challenges for the expert reviewer(s). Their quality depends heavily on the expertise and vested interests of the evaluator(s), the theoretical framework that informs the evaluation and the time allocated to and investment in evaluation. Policy-makers might consider the benefits of a panel-type approach for full programme evaluations, where a heterogeneous panel is composed not just of evaluators from the same donor-oriented consulting background or agency, but the use of independent expert evaluators who bring to the task a range of different specialist skills and North-South perspectives. This might include, for example, academic experts, evaluation experts (skilled in specific methodologies for measuring impact or conducting large-scale surveys), academic monitoring and management experts, user-end beneficiaries. Another approach would be to separate out different types of evaluations, using different methodologies and evaluators. The sole use of small focused user-reviews and impact studies in the sample, for example, provide new perspectives and detail on particular aspects of an intervention that would otherwise be missing in an expert review.

Donors funding these programmes are making extensive use of evaluation frameworks, e.g. log frames and OECD development criteria often accompanied by a set of indicators. These clearly provide a strong frame of reference for the evaluator, but there is an accompanying risk of a tick-box approach to measurement and for evaluators to be straitjacketed by the provided indicators where more flexibility might capture unintended outcomes or open up the focus of evaluation, particularly to reflect contextual successes or reaching specific human skills calls.

The EdQual program⁴⁰, in particular, appears to underline the value of very careful planning for both research programme design and its subsequent evaluation based on a prior examination of the strengths and weaknesses that have emerged in earlier similar projects, e.g. implementation problems due to the tendency for partnerships to be dominated by northern

⁴⁰ <http://www.edqual.org/>

partners to the extent that they fail to achieve mutual learning, shared objectives and joint achievements. This seems to have been avoided by carefully planned decentralised African-led programmes and evaluation that involves the beneficiaries as equal partners.

A key issue in these type of programmes and their evaluations is whether there is successful evidence of a sense of ownership, transferral of leadership and independence (e.g. beneficiary-defined and initiated projects). The training for and use of local level evaluation (beyond stakeholder consultation) plus an openness to new types of methodologies seems a vital but overlooked ‘ownership’ ingredient and which, of course, in itself contributes to research capacity building.

Institutional Support

Much the same points made about methodologies used in links and consortia programmes apply to those for institutional support.

Main Gaps

1. Summarising here some of the gaps in development evaluation, there is a general argument for more appropriate methodologies to meet the growing evaluative demands from donors and beneficiaries, and as a consequence of larger, more complex, multidimensional interventions that are proliferating.
2. Nuanced insights into the effects and impact of donor interventions is hampered by a shortage of high quality, sophisticated evaluations which employ a range of methodologies, which mix both qualitative and quantitative methods and are built into the design of programmes so that more long-term evaluative procedures are in place rather than retrospectively added. The concept of ownership in evaluation, a balance of North-South perspectives and interests, is conspicuous by its absence.
3. The growing demand for impact evaluation as the basis for evidence-based policy in HE development is limited by a lack of focus on issues of attribution – tracing out cause and effect – and quantification (e.g. instrumental variable testing, propensity scoring, regression discontinuity and difference-in-difference). There is a shortage of rigorous, quantitative assessment which through regression analysis, or some other appropriate methodology, tries to get at the long-term. It is equally important, though, to balance these approaches with

Implications For RUFORUM

a concern for context, criteria and the limitations of evidence-based quantitative evaluations. The drive for results has placed an over-emphasis on the products of an intervention rather than appreciation of the processes and necessary stages that interventions go through to reach products. The issue here is that funders tend to see the objective of education as simple and measurable, whereas education specialists know it to be more complex. One strong tradition of curriculum development argues for concentrating on process and against reducing the educational process to neatly measurable objectives (e.g. Stenhouse⁴¹, 1975).

4. Cost studies are a much overlooked area and, like impact studies, could only be addressed by planned training for and use of a range of approaches (cost analysis, cost effectiveness, cost benefit). Again, the reliability and quality of monitoring systems plays a vital role here.
5. One clear gap is the lack of well-managed monitoring systems at both the programme and project level from which evaluation draw their baseline data.

ALINe has been working to support RUFORUM with its work on Monitoring, Evaluation and Learning since 2010, using the Theory of Change as a central tool to aid the process. This has enabled staff at the Secretariat to reflect on their existing understandings of how change happens as a result of their own interventions as well as the influence of other actors and factors that may be beyond the Secretariat's direct control. It also helped staff to arrive at and organise a set of relevant indicators for tracking the different kinds of changes that RUFORUM seeks to contribute to at various levels: within universities/faculties, at the level of RUFORUM as a dynamic regional consortium, and at the level of the wider agricultural sector (national governments, employers, users of research outputs, etc.).

The initial TOC has since been opened up to further discussion with other members of RUFORUM, including Deans and PIs amongst others, which has helped to enrich the TOC by contributing insights into change processes, dynamics, opportunities and constraints at different levels. These

⁴¹ Stenhouse, L. 1975 *An introduction to curriculum research and development*, London: Heinemann

discussions, inasmuch as they offer new insights, also raise new questions to be answered. As a result the TOC functions as an evolving record of current understanding of how change happens. However, working with an evolving TOC involves managing the tension between establishing routine mechanisms of data collection (process and outcome monitoring) and ongoing changes in what is to be measured. This is addressed by defining certain core elements to be monitored for a certain period regardless of shorter-term changes in the TOC.

The monitoring of process and outcomes through the TOC is expected to generate a body of data that will enable RUFORUM to test whether or not the desired changes are taking place, to what extent this is happening and what key stakeholders think about these changes, including their satisfaction with the RUFORUM Secretariat's role in providing the necessary support. By bringing this evidence together and analysing it alongside the RUFORUM TOC with the involvement of relevant stakeholders, it is expected that an evidence-based discussion about the changes to which RUFORUM is contributing will be possible. This will at once produce data that can be used to take stock of and report on progress (e.g. for upward accountability purposes) but it will, perhaps more critically, permit a more nuanced discussion about the TOC, about how change happens, about challenges and opportunities and how best to address them.

At the same time, the TOC and the selection of indicators to measure change reflects certain limitations. M&E consumes resources and, simplifying somewhat, the more comprehensive and rigorous the methods used, the more costly it is. Thus RUFORUM faces, for example, difficult questions: deciding to what extent it wants to track changes at the farmer level, what mechanisms, if any, it wants to use for consulting with farmers on the TOC and the results of evaluations, whether and how to seek feedback from employers of the graduates it has trained and to what extent it wants to track its influence on policy processes. Each of these choices raises methodological challenges, closely coupled to resource considerations, particularly if a measure of impact (defined here as evaluation that uses a counterfactual) is sought. Trade-offs must therefore be made between certain conceptions of rigour and validity (in which the clinical trial typically serves as gold standard) against considerations of utility and feasibility. The prioritisation of the latter, particularly where the primary goal is change in a complex

system, typically means an increased concern with keeping track of trends and ongoing changes both within and external to the initiative, using a mixture of methods (qualitative, quantitative), tracking different changes to differing degrees, regularly reflecting on and refining the indicators, tools and methods (including the TOC) used to identify issues, gaps and challenges as they emerge and to respond to them in a manner that takes on board the perspective of a wide range of relevant stakeholders. In this sense, the role of M&E can be, perhaps, better understood as generating an evolving and tentative model of the system to be changed and then tracking changes in that system through various forms of empirical verification.

Critically, this does not mean a neglect of concern for changes that are beyond the immediate influence of RUFORUM. Rather, it raises the question of which actors are best placed to make rigorous assessments of the changes that are taking place at the boundary of what RUFORUM can realistically expect to keep track of. It thus raises questions in particular about the roles of universities/faculties in tracking and evaluating various changes themselves, whether this concerns the impact of their projects and programmes and tracking what has happened to their graduates and research outputs. It also raises questions about the role of national governments and related (statistical, sector-wide) agencies who may be better placed to make assessments of wider shifts in the agricultural sector and, in particular, the contributions of universities at this level.

Conclusions

The main purpose of this study was to examine literature and evidence to understand whether and how HE interventions for development were being evaluated. The work also reflected on the evidence for impact, how and where it has been attempted.

The following broad observations on effectiveness of interventions include:

Education and Training

- a. There are good successful completion rates for many programmes.
- b. Evidence on brain drain is mixed (and needs to be interpreted in light of recent and nuanced discussion of this issue).
- c. There is confirmation that rigorous and equitable selection methods, and guidance for selection, have worked (as implied in (a)).

- d. Scholarship programmes have benefited in various ways (e.g. good completion, happier learning experiences, less brain drain) where both sending and receiving countries have roles to play in the selection process.
- e. There is evidence from various evaluations of (1) alumni applying what they have learned (2) participation in the scholarly community (3) wider effects through their participating in the community (4) training others (e.g. in agriculture and supervising PhDs).

Links and Consortia

- a. Successful mobilising of new knowledge resources from other sources (e.g. UK HEIs, African one-stop research sites);
- b. Significant transfer of research knowledge and skills to developing countries;
- c. Evidence of promoting innovative initiatives in MDG areas that evolved into broader programmes and community extension activities;
- d. Goodwill and long-term links between northern HE institutions and partner southern HE institutions, and also for South-South HE institution partnerships. Growing evidence of shift towards home grown, new local level courses, doctoral programmes, leadership and skills in competitive funding proposals and towards the establishment of centres of excellence;
- e. There is evidence from various evaluations of (1) beneficiaries applying what they have learned (2) participation in the scholarly community (3) wider effects through their participating in the community.

Institutional Support

- a. Evidence of successful capacity building in policy, infrastructure and academic support systems (e.g. finance, QA, ICT, library) in targeted HE institutions and arising from long-term approaches;
- b. Evidence of institutional benefits in the form of recognition of the value of research in universities, organisation benefits in the form of improved ability to build on research, and financial benefits in terms of continued investment in research and financial incentives to researchers;
- c. Strong evidence of North-South and South-South collaboration raising capacity in research and academic support systems;
- d. Direct positive contribution to the scientific knowledge base;
- e. Basic university frameworks in place for regional harmonisation (QA) & networking;
- f. Expansion in indigenous doctoral programmes.

The following observations have been distilled from the study and the conference:

1. There is general consensus that HE investment is beneficial to the economy and society but its development value, e.g. in reducing poverty in LMICs, is less clear. Historically, investment has served many objectives, not just development.
2. But there appears to be little formal evidence for the development benefits of the three kinds of HE investment mentioned. Further, there are potential negative development effects of HE interventions, including ‘brain drain’ and the continuation of inequitable, Northern-driven research partnerships. Finally, the past decade has seen dramatic changes in HE: the internationalisation of universities, developments in Information and Communication Technologies, and the rise of private HE provision, particularly in LMICs, all of which may have changed the landscape for HE interventions for development.
3. Where development is an objective, HE investment lacks a convincing ‘theory of change’, particularly in how investments at the individual, organisational or institutional level will lead to development outcomes and positive societal impacts. The relationship between investment in individuals and effects at the institutional and societal level is particularly poorly understood.
4. Evaluation of HE programmes has focused very strongly on outputs, e.g. individuals trained to a certain standard, and very little on outcomes, in terms of measurable development contributions from individuals or institutions, and even less on societal impacts. There has been more evaluation of training and scholarship programmes than of other kinds of interventions.
5. As with much evaluation of complex interventions, there has been little comparison of evaluations of HE programmes done in different countries, contexts and sectors (e.g. agriculture, health), nor has there been much effort to compare the value of different approaches.
6. Evaluation of HE for development has been strongly Northern-driven, associated with donor views on ‘value for

money', whereas Southern perspectives on 'what works' are rarely sought.

7. For all these limitations, there are new and innovative tools being developed and tested for evaluating HE interventions, and considerable potential to adapt and apply a range of new methods from other development areas. These can be built into new interventions from their outset, or used to analyse historical data sets provided by long-running programmes.
8. In order to strengthen the evaluation of HE interventions for development, discussions at the LIDC conference suggested the following actions:
 - a. First, we should extend and complete analysis of different kinds of interventions, their objectives and their evaluation, with the specific aim of bringing together experiences and tools across interventions and sectors as a resource for further evaluation effort. This extended study should incorporate evaluation methods and experiences from programmes in transition and high income countries, so as to capture all useful information.
 - b. Second, there is a need to develop and test theories of change for HE interventions that address ways in which individual, organisational and institutional interventions will have development outcomes and impacts. This would draw upon evidence collected in the extended study above, and should involve participation of Southern HE institution leaders and policy makers. Specific projects which might be used to do this include:
 - i. analyses of historical data sets (e.g. scholarship programmes) in order to identify outcomes and validating theories of change
 - ii. projects based in LMICs which examine at an institutional or national scale the outcomes and impacts of multiple, past HE interventions.
 - c. Finally, we should design new impact evaluation tools for inclusion in new HE interventions, building on current experimentation with real-time evaluation, and experimenting with a range of quantitative and qualitative methods and the use of counterfactuals.
9. The conference brought together a community of development, education and impact evaluation specialists and demonstrated a common interest to advance this area of

study. It was, however, comprised largely of development organisations and Northern HE institutions and further efforts should engage more fully governments and universities in LMICs.

All capacity-development interventions are based on theories of some sort. However, they tend to be theories that are implicit in the minds of those who design and implement the interventions, rather than explicit theories in the form of coherent narratives that can be discussed, debated, improved and shared. Because there is seldom a consensus among key stakeholders on the programme theory, various actors may have different – sometimes conflicting – concepts of the programme's goals and strategies and how its activities are expected to strengthen capacity.

ALINE's literature review has highlighted the relevance of individual approaches to capacity development within broader debates in the capacity development literature. It has put forward the key elements and indicators to consider in an evaluation of a fellowship, including a reflection on the different methods available to assess these programmes. Lastly it has explored the literature on agricultural research and development and impact evaluation to find alternative and complementary participatory methods to incorporate notions about complexity and the multi-stakeholder nature of research, and its impacts on social change.

In conclusion there are six lessons learnt from the literature review:

- The importance of process;
- The need to focus on learning;
- The benefits of using several methods;
- The need to include stakeholders in inclusive participatory ways;
- The importance of initiating processes of reflection about assumptions and power; and
- The need for ARD programmes to understand that different kinds of technologies have different social outcomes and that a technological solution may not be the desired solution every time.

The Importance of Process

When designing an evaluation for a fellowship programme, emphasis should not be on research outputs such as an evaluation report. The learning that can be sparked through

Pinto, Y.

the process of evaluation itself could be significantly more useful than with the ulterior use of the evaluation outputs. This is particularly true when the evaluation involves different actors within the organisation – managers, M&E staff, project managers, mentors, fellows, etc. – and when it engages with different stakeholders. Thus an inclusive approach that fosters debates both internally and externally has more potential to generate reflexivity and organisational learning.

As indicated by Horton and Mackay, participation and process can have lasting effects on knowledge, attitudes and skills of people, and ultimately affect their behaviour and decisions. The question that an evaluating team must raise is: do we have the time and resources required? Programmes must have budgeted time and resources to engage with stakeholders for ongoing evaluative processes (theories of change workshops, deliberation on projects, discussion on impacts, etc.). This means that when designing the evaluation, one must reflect on who needs to be involved, when and how to ensure ownership, inclusion and empowerment as well as identifying who will be the facilitator in this learning process and how they will be selected.

Another question that arises is who will benefit from the lessons learned in the evaluation, and how will the lessons learned be used? It is important not to take a ‘gap’ approach to the organisations involved (in this case the fellowship programme and its partners such as the African research centres), but to focus on those existing endogenous capabilities. Thus this assessment has to contribute to and catalyse these endogenous processes for change; the evaluation must tap into existing energies. Involving partners in the design is a way of ensuring we build on these energies.

A Focus On Learning

The lessons learned from evaluations must translate into future change. Will the fellowship programme continue? Will it work with the same partners? If the programme staff and organisations do not continue working together, the lessons learned become less useful, and merely acquired for accountability purposes. If the programme does continue, maintaining in some degree the same institutions and people albeit in a different shape, there are more pragmatic lessons to be learned and therefore lasting change is more likely. For evaluation results to be useful for behavioural change, it is fundamental to look toward the future in practical ways.

Evaluations need the space to be critically reflective. A strong focus on accountability – for instance, if the evaluation has been demanded by the donor to assess performance and therefore future funds depend upon it – will inevitably lead to a ‘culture of success’, in which different actors are pressured to give evidence of good practice. Delinking M&E from funding will yield more effective learning.

If delinking is not possible, the donor involved must at least be aware that short-term projects and a ‘culture of success’ often deters organisations from being self-critical and from engaging in organisational learning that tackles structural issues. They therefore may not be able to produce tangible benefits in a short period of time. Commitment should thus be long term and expectations regarding tangible results should be flexible – their involvement in the process should be in a partnership form.

When assessing the impacts of a specific capacity development programme such as a fellowship, it is also important to be flexible in its requirements of attribution. Social change in ARD is complex and multiple capacity development providers work with different organisations. This literature review has shown that attributing a particular impact to a particular donor-funded scheme is quite problematic. There are two ways forward: either engaging in sector-wide impact assessments conducted by sector platforms, evaluating the impact of the broader system in which research is embedded; and/or delivering evidence of contribution and progress that is not strictly attribution, for instance, illustrations of change, recording changes that have occurred, and highlighting specific examples that illustrate wider changes.

Using Multi-Methods

In order to evaluate an ARD fellowship, the best option is to combine several methods.

Questionnaire

An evaluation of a fellowship often includes a questionnaire for project managers, supervisors in home and host research centres, mentors and fellows including the different variables outlined in Section 2, regarding design, effectiveness and impact. A questionnaire should include questions on:

- The coherence and relevance of the programme
- Each component and how they contribute to programme goals

Design

Effectiveness And Implementation

- The appropriateness of candidate and research project selection
- The appropriateness of programme duration
- The appropriateness of programme location
- The degree to which the organisational development of home organisations was integrated into the design
- The degree of involvement with stakeholders

- The running of the programme
- The degree of implementation of activities and results obtained
- The quality, motivation, time and resources allocated for the programme managing team
- On whether and to what extent the expectations of fellows and supervisors in home and host centres were met
- The degree of learning acquired by fellows
- The subsequent changes in behaviour
- The development of soft skills and personal development (self-awareness, leadership and management skills, confidence and motivation, capacity to become role models, etc.)
- The contribution of the programme to organisational capacity (in both host and home organisations)
- Fellow mobility and causes (if applicable) of brain drain
- The quality and impact of mentor relations
- The effect on networking
- The impact on partnership building
- Cost-effectiveness (to project managers)

- Impact on fellows' careers: publications, promotions, etc.
- Impact on fellows' behaviour
- Projects taken up by the fellow after the fellowship
- Fellows' perceptions of the impact of their work on technology adoption and changes in production.

The evaluation team will need to ascertain if it is relevant to launch a career tracker, surveying fellows every year (or more frequently) to see how their careers unfold. This is most relevant if the fellowship programme is likely to continue for several years.

Interviews

To gain a more nuanced understanding of the issues above, a random selection of fellows, supervisors (home and host) and mentors should be interviewed, as well as the programme management team and M&E officers, and a selection of

CV Analysis

relevant stakeholders. These interviews should be in depth (from 60 to 90 minutes) and should be semi-structured, assessing the critical decision points in a fellow's career and the dynamics of programme implementation.

Participatory Impact Exercise

To assess the impact on fellows' careers, a search should be conducted on publications, citations, projects designed, and funding grants (those applied for and those successfully obtained).

Stakeholder Deliberation

In order to gauge the impact of the programme, it is interesting to organise a participatory exercise such as the Most Significant Change (MSC). In this method, the most important thing is the discussions at different levels to identify the MSCs and not the resulting MSC itself. What adds value and representativity is the quality and inclusiveness of the participatory processes to identify the MSCs at the different levels.

Depending on the resources, it would be interesting to bring together the different stakeholders involved, in a workshop or a series of focus groups, to discuss the project design, implementation and impact, raising some of the questions indicated above.

This stakeholder deliberation is particularly useful to appraise the research projects carried out by fellows and their relevance for small-scale farmers. Similarly, it is a useful way to question stakeholders about the overall impact of the programme on changing small-scale farmers' livelihoods. It also helps in identifying and learning about the 'chain of events' that leads (or not) from a capacity building exercise to technology adoption, to livelihood changes.

It is also useful to have a workshop on theories of change (or an impact pathways analysis), to explore in detail the assumptions that different actors have about this 'chain of events' and how that affects the impact and relevance of these programmes. This exercise will be most useful if the scheme continues in the future in some form or another, so that practical lessons learnt can be carried over.

Stakeholder Participation

The ALINE literature review strongly emphasises the importance of stakeholder participation at all stages of the capacity development cycle, from design to evaluation. In the case of a fellowship evaluation related to ARD, stakeholders

should include not only those involved directly in the fellowship – donors, programme managers, fellows, supervisors (host and home research centres) and mentors – but also those indirectly involved in broader ARD processes. These include the ‘broad system of agricultural research’, such as universities, policymakers, extension workers farmers, farmers’ unions, NGOs, etc. Who the stakeholder participants are will depend on the objectives of the evaluation and the conditions required to give enough weight to the end-users’ voices.

In ARD fellowship programmes such as the AFP, the end-users of technology or research outputs are very often small-scale farmers. Ultimately fellowship programmes and ARD aim to change the livelihoods of poor farmers. Capacity development literature shows that user-catered capacity development evaluations lead to improved learning outcomes. Small-scale farmers must have a voice in the review of performance of agricultural research, extension and training systems. In this case, these users should have a voice in the evaluation design and should have a central role in the appraisal of research projects and in the evaluation of the programme as a whole. As discussed in the literature review, great care has to be taken not to interpret the category ‘farmer’ or ‘small-scale farmer’ as an homogeneous one: a farmers’ association does not represent necessarily all farmers; technology will influence different kinds of farmers differently depending on context (for example, small or large-scale farmers, rainfed or irrigation agriculture, marginal areas, agricultural labourers, tenants or landowners, etc.). A stakeholder discussion must acknowledge the complexities of rural social dynamics.

It is important to note that it is not required for multi-stakeholder consultations to reach a consensus. It is only necessary to have some sense of shared understanding and values as a starting point for the capacity development assessment. Consensus should not be an objective (and in any case it is not a realistic objective). Evaluations should explore the diversity of visions and opinions to enrich the learning process, always making sure that the most powerful actors (donors, project managers, etc.) do not co-opt (voluntarily or not) the process and instead do give space for marginalised voices to speak.

Exposing Assumptions and Power Relations

Are we being critically reflective and are we prepared to challenge our assumptions? Capacity development programmes should encourage stakeholders to reflect on how each has

different notions about how change occurs, and these assumptions are often implicit in stakeholder actions. Different development actors involved in an individual approach to a capacity development programme will have different theories of change and will thus act according to their assumptions, shaping the development outcomes. Similarly, a divergence between different actors' premises will generate conflict and contradiction.

A theory of change exercise (or a very similar process such as the impact pathways analysis) is a good way to put assumptions at the centre of analysis and recognise the non-linear, complicated or complex nature of development. A fellowship evaluation should have this type of element to make assumptions explicit. For example, AWARD has included a workshop of this kind in their work.

Increased capacity of scientists—research outputs—emergence of new technology—adoption of new technology by small-scale farmers—increase in agricultural productivity—reduction in poverty.

Different stakeholders then discuss the different pathways that (may) link the capacity building activities of the fellowship to the ultimate changes in small-scale farmers' livelihoods. The more the exercise allows for diversity of opinions and debate, the more it will become a learning exercise and the less it will become a stepping stone to obtain a convoluted log frame (making it therefore less useful).

It is important to note that the theory of change exercise in a fellowship should strive to explore the assumptions beyond the area of control of the programme – from capacity building to fellows producing a certain research output. It should look into how fellows may organise other projects/initiatives, the research outputs transformed into technologies, how small-scale farmers adopt those technologies, and what happens to farmers who change their practices and those who don't.

Questioning must also cover broader programme assumptions such as the location and nature of support included in the scholarship. For example, some actors may see a value in investing in developing country research centres and paying fellows to remain in those centres; others may think fellows will be able to bring back more capacity if their training occurs in a northern research centre.

‘Good Science’ is Insufficient

Lastly, stakeholder deliberations must reflect on power relations: individual approaches to capacity development take place in an environment that is messy and political, where certain types of knowledge are privileged over others, where the voice of the marginalised is silenced. Thus capacity development initiatives should strive to expose these power differentials and aim to redress them encouraging real partnership. We have to acknowledge the role of hierarchy and power in the partnerships formed to deliver these programmes. Deliberation must focus on what knowledge is prioritised and what power relations emerge between partners such as donors, northern NGOs, southern NGOs, and southern research centres. Ultimately we must ask: are we seeing donor-driven capacity development and, if we are, how is this curtailing or facilitating learning?

Research fellowships (as opposed to taught fellowships) often entail the selection of a candidate and a research project. The relevance and quality of these projects must be assessed both in the implementation of the scheme and in its evaluation.

It is particularly important to note that candidates’ research projects should not only be evaluated with regards to their scientific quality (a necessary condition), but also with regards to their relevance vis-a-vis the programme objectives. An assessment panel (which should include not only agricultural scientists but also social and development scientists) should assess the real use of those potential research outputs in the lives of poor small-scale farmers, ensuring not only that the science behind the research is good, but that the potential technologies are relevant for the models of production used by marginalised small-scale farmers.

It is also important to integrate into the programmes the realisation that technology and research outputs are not a silver bullet in development processes. Technological change may be appropriate in some contexts, but not in others. Often the livelihoods of small-scale farmers can be improved through other means – for instance, economic or legal changes, or new social arrangements. ARD must engage with broader development processes, acknowledging that improved agricultural technologies are only one piece of the development puzzle.

Appendix 1: Challenges And Methodological Approaches

Non-linearity and complexity. Impact assessments include a vision of technical change that assumes that causes can be separated and analysed individually. This gives an emphasis on

prediction, measurement and control⁴². Yet complexity theory has indicated that in technological change multiple, unpredictable factors are at play⁴³. Because of the complex nature of social change it is not possible to predict the relation between cause and effect, and our interventions will need to understand and adapt to that complexity. Complex systems are path-dependent and highly unpredictable. In Eyben *et al.*'s⁴⁴ words:

Complexity theory posits that change is emergent. Organised efforts to direct change confront the impossibility of our ever having a total understanding of all the sets of societal relationships that generate change and are in constant flux. New inter-relational processes are constantly being generated, which in turn may affect and change those already existing. Small actions may have a major impact, and big ones may have very little impact.

In reality alternative ‘pathways’ from capacity development activities to outcomes and impact are not linear but branching. Capacity development activities will depend on the ‘trajectory and momentum’ already existent but they will also depend on other factors – actors’ interests, policy influences, power structures, culture, weather, etc. that interplay with these processes. Ortiz and Taylor (2008) see in a successful M&E approach to capacity development in a complexity system as one that embraces this complexity. This constant reassessment is called ‘iterative measurement’².

Assumptions about change. If social change is uncertain, complex and emergent then actions that aim to promote particular forms of social change i.e. developing the capacity of individual researchers so as to eventually improve farmers’ livelihoods, are based on particular premises, assumptions and value that underpin the change effort⁴⁵. Further, different stakeholders involved in the process from the design and implementation of the programme to its evaluation will have different notions on how change occurs and will act according

⁴² Ekboir, J. (2003) ‘Why impact analysis should not be used for research evaluation and what the alternatives are.’ *Agricultural Systems* 78: 166-184.

⁴³ Ortiz, A. and Taylor, P. (2008) *Emerging Patterns in the Capacity Development Puzzle. Why, what and when to measure?*: p14, Brighton: IDS.

⁴⁴ Eyben, R., T. Kidder, et al. (2008) ‘Thinking about change for development practice: A case study for Oxfam GB’, *Development in Practice* 18.12.

⁴⁵ Aragón, A. O. (2009) ‘Interpreting Worldviews and Theories of Change on Capacity Development of Social Change Organisations’: p13, PPSC Team’, Brighton, IDS

to their assumptions, shaping the development outcomes. Theories of change workshops make explicit and share with others the assumptions about how change happens. Acknowledging divergence in assumptions, premises and values between stakeholders and the nature of the relationships between activities, outcomes and impacts is crucial. Engaging stakeholders in assessments is valuable in identifying the assumptions about the pathway and potential pathways for collaboration emerge. Active stakeholder participation in evaluations is a form of capacity development in itself. It can lead to ‘lasting effects on the knowledge, attitudes and skills of people and on their subsequent decisions and actions’ (Horton and Mackay 2003: 133)⁴⁶.

Multi-stakeholder. Not only do we need to take into account the fact that different actors negotiate spaces for change according to their own values, interests and shifting development outcomes; but also that they will have very different ideas about what impact is. Some actors may value economic notions of productivity, while others may value different development results such as empowerment or autonomy.

The problem of attribution. Capacity development schemes that work with individuals or organisations encounter difficulty in proving to others (particularly donors) their role in achieving certain outcomes. This is because of the complexity of social change, but also because there are multiple capacity development providers working with the same organisations. There are multiple examples of capacity building programmes funded by many donors and it is difficult to attribute impact to differing contributing schemes. This can be solved by carrying out organisational assessments of schemes instead of evaluations⁴⁷ so as to see the effect of all different capacity development programmes in their performance.

Engaging with sector platforms (e.g. RUFORUM) both in strategy building and impact assessments is an interesting move for individual approaches to capacity development schemes. Impact assessments should focus not only on research inputs and outputs, but on the ‘broader system in which research is

⁴⁶ Horton, D. and R. Mackay (2003). ‘Using Evaluation to Enhance Institutional Learning and Change: Recent experiences with agricultural research and development’, *Agricultural Systems* 78: 127-142.

⁴⁷ Simister and Smith 2010.

embedded and the contribution of research to technical and institutional changes' (Watts, Horton *et al.*, 2008: 23)⁴⁸. This is because innovation involves not only research tasks but also the influence of non-research actors such as the extension services, the relevant government bodies, university management, the agricultural NGO sector, farmers' unions and many others. Researchers can contribute most effectively to innovation when they participate in partnerships, alliances and coalitions.

Engaging researchers with end-users in these sector-wide platforms is a useful learning exercise as often researchers possess little knowledge of how innovation occurs in agriculture. They do not attempt to understand how research and development (R&D) can create change at the farm level, nor the mechanism that may link technology to productivity and broader societal benefits⁴⁹.

Yet difficulty in achieving attribution should not be over-emphasised, as long as there is evidence of contribution and progress. When capacity building – as in the case of fellowships – is concerned with altering the vision, values and culture of researchers and their organisations, or aims to improve organisational systems such as research planning, fundraising or human resources, it is difficult to trace all wider results as they spread over time and space.

The opposing forces of accountability and learning. Donor pressures to make capacity development programmes accountable have had the unintended effect of decreasing the potential of capacity development to promote organisational learning. Impact assessments and evaluations are fundamentally directed at gathering information for fund-raising, and satisfying donor requirements.

According to Horton and Mackay, most impact assessments based on economic valuations of initiatives have included accountability requirements, yet these have not helped 'managers and scientists understand how their activities produce results or to draw lessons about how to improve future research

⁴⁸ Watts, J., Horton, D. *et al.* (2008) 'Transforming Impact Assessment: beginning the quiet revolution of institutional learning and change', *Journal of Experimental Agriculture* 44: 21-35.

⁴⁹ Springer-Heinze, A., Hartwich, F. *et al.* (2003) 'Impact pathway analysis: an approach to strengthening the impact orientation of agricultural research.' *Agricultural Systems* 78: 267-285.

programmes and manage the organisations that undertake them' (Horton and Mackay 2003: 132).

To promote institutional learning, impact assessments must link assessments more closely with management decision-making.

a) Types of impact assessment and potential applications for evaluations of capacity development initiatives

Building on the debates about complexity, different organisations have:

- Focused on mostly quantitative experimental economic analyses that aim to 'pin down' complexity in their modelling
- Have adopted a qualitative approach integrating complexity, such as the Impact Pathway Analysis
- Have adopted a collaborative approach to impact assessment, seeing knowledge as co-constructed with stakeholders.

These different approaches are discussed below.

(i) Experimental economic impact assessments. They heavily quantitative assessments are often carried out by agricultural economists, and are mostly based on the projected economic impact on poverty. The CGIAR and the World Bank have traditionally favoured this type of impact assessment. They are based firstly on microeconomic analyses to measure the impact of adoption on individual adopters (before diffusion occurs), and then on measurements of aggregate impact of a continuously evolving line of various improvements⁵⁰. To compare between adopters and non-adopters (the counterfactual), the studies are based on randomisation at a village level (to integrate spillovers), where some villages have adopted the technology and others have not. Similarly another variable is early stages of adoption and later stages of diffusion⁵¹.

These experimental tools are not necessarily appropriate for impact evaluation of individual approaches to capacity development in ARDs.

⁵⁰ de Janvry, A., A. Dunstan, et al. (2011) 'Recent Advances in Impact Analysis Methods for Ex-post Impact Assessments of Agricultural Technology: Options for the CGIAR', Berkeley: University of California at Berkeley.

⁵¹ Baker, J. (2000) 'Evaluating the Impact of Development Projects on Poverty: A Handbook for Practitioners', Washington D.C.: World Bank.

(ii) Impact Pathways Analysis: a qualitative approach to impact assessment. Impact Pathways Analysis is a very similar exercise to that carried out to determine the organisations' theories of change. It lays out the logical links from research inputs to economic, environmental and social changes, enquiring about external processes that may influence different stages in the pathways. It breaks the pathway down into smaller sequences of events, adding intermediate steps. The Impact Pathways Analysis is rarely used alone; it is a complementary technique. For example, the CGIAR advocates its use together with an array of quantitative approaches (including cost-benefit analyses and the economic impact) as well as other participatory approaches such as outcome mapping⁵². Its usefulness is that among the different theories of change; it exposes and promotes debates about the assumptions, values and premises that occur regarding social change and development agents' role in trying to transform agricultural systems. As mentioned above, a process like this must integrate the diversity of stakeholder worldviews and opinions about the project and generate debate. It has to be an iterative process, as different external and internal factors will arise during the project and beyond it.

(iii) Collaborative participatory approaches. These methods are based on an understanding of knowledge as co-constructed, and engage researchers, users and other stakeholders in designing, monitoring and evaluation programmes. They can be incredibly useful when evaluating individual approaches to capacity development schemes such as fellowships.

b) Outcome mapping

More and more organisations are showing interest in outcome mapping, from donors to capacity development providers and consultants. For those organisations that have used it for planning purposes, it is a tool that allows stakeholders to assess the contribution of a development project or capacity development scheme to specific outcomes. As it involves stakeholders in a participatory fashion, it is also a learning tool. Developed by the IDRC, its innovative approach focuses on behavioural change or the 'people' side of social change.

⁵² Walker, T., Maredia, M., et al. (2008) 'Strategic Guidance for Ex Post Impact Assessment of Agricultural Research. S. P. o. I. Assessment', Rome: GGIAR Science Council.

Outcomes are defined as ‘changes in the behaviour, relationships, activities or actions of the people, groups and organisations with whom a programme works directly’. These behaviour changes are logically linked to programme activities, yet it acknowledges that outcomes need not directly be caused by these activities. The starting point is the desired outcome in terms of changed behaviours, and then it works back identifying those factors that may lead to that outcome, including the project activities. This means this method focuses less on attribution to outcomes and more on contributions. It is good at establishing change at an organisational level¹⁶.

It then goes on to identify boundary partners – actors that are likely to influence the outcome – and thus tries to engage and collaborate with them. It is a good tool to identify opportunities for influence. When evaluating, it will be possible to see if boundary partners were in fact influential, and to gauge the relative success of capacity development exercises in mobilising their support⁵³.

To use outcome mapping effectively, it needs to have been used since the planning stage, so as to contrast the achievement of desired outcomes, the behaviour change of boundary partners, and stakeholders’ perception of the ‘reality’ that occurred. The strength of this tool is that in its three stages it includes design, outcome and performance monitoring and evaluation. The whole exercise is specific to the organisation and context; and indicators are identified specifically for each organisation⁵⁴.

c) Stories of change

These participatory tools emphasise that social change cannot be illustrated well by only a set of indicators but that there is a need to understand the depth and complexity of change. These methods involve a specific narrative of change at an individual, organisational or societal level⁵⁵.

⁵³ ODI (2011) ‘Outcome Mapping’, Research and Policy in Development: Toolkits from www.odi.org.uk/rapid/tools/toolkits/Communication/Outcome_mapping.html.

⁵⁴ Earl, S., F. Carden, et al. (2001) ‘Outcome Mapping: Building learning and reflection into development programs’, Toronto: International Development Research Centre.

⁵⁵ Soal, S. (2001) ‘Measuring Development. Holding infinity’, from Soal, Sue.

d) Most Significant Change (MSC)

The most widely used stories of change tool is the approach known as the Most Significant Change (MSC). This enables monitoring without indicators and evaluations to take place. It is particularly useful in cases where it is not possible to predict exactly what the outcomes will be; where outcomes vary depending on the beneficiary; where there is divergence of opinion between stakeholders on which outcomes are the most important; and where interventions are expected to be highly participatory, including any forms of monitoring and evaluation of the results⁵⁶. These are all applicable to individual approaches to capacity development, provided they use participatory approaches.

MSC aims to identify significant changes brought about by a development intervention, especially in those areas where changes are qualitative and therefore not susceptible to statistical treatment. It relies on people at all stages of a project or programme meeting to identify what they consider to be the most significant changes within pre-defined areas (or domains) (Simister and Smith, 2010: 16).

The MSC process collects stories of significant changes from field level, and these are systematically selected by panels of stakeholders or programme staff. Once these changes have been gathered, workshops including different stakeholders read the stories and discuss the value of the reported changes to decide which is the most significant of all. There may be multiple levels at which stories of changes are gathered and selected⁵⁷.

e) Peer Ethnographic Evaluation and Research (PEER)

This method is also reliant on the richness of narrative data to illustrate the complexity and particularities of social change. The key idea is to train members of a community – peer researchers – to carry out in-depth conversational interviews with colleagues in their social networks, and to report back findings in a series of in-depth interviews with a social scientist. It builds upon the creation of trust between researchers and the community. Peers are often recruited among marginalised

⁵⁶ M&E News (2011) Most Significant Change (MSC). What is MSC? from <http://mande.co.uk/special-issues/most-significant-change-msc/>.

⁵⁷ Davies, R. and J. Dart (2005) ‘The ‘Most Significant Change’ Technique: A guide to its use’, Care International.

and harder to reach social groups. Through getting insider knowledge PEER collects local interpretations of social change and can attribute these to development interventions and other external factors⁵⁸.

f) Appreciative inquiry

Originally designed to improve organisational performance in private companies, appreciative inquiry is a project design and M&E tool based on focusing on ‘strengths’ rather than ‘gaps’ or ‘weaknesses’. It explores the ‘best of what has been’ i.e. it is appreciative, and it aims to be applicable, provocative and collaborative⁵⁹. It aims to create a collective image of a better future in the organisation based on ‘what works’. The process aims to discover best examples of organising and organisations within the institutions, to understand what forces – the people, the organisation or the context – create better performance, and finally to reinforce and amplify the people and processes exemplifying best practice. The aim is to build a momentum of good practice within an organisation.

The process is very similar to MSC in its bottom-up approach of collecting stories of good change within an organisation. It involves:

... a bottom-up interview process where almost all organisational members were interviewed to uncover the “life-giving forces” in the organisation. People were asked to recall times they felt “most alive, most vital, most energised at work” and were then questioned about those incidents. The interview data were then treated much the same as any qualitative data set; through content analysis the consultants looked for what people in the organisation valued and what conditions led to superior performance. This analysis was fed back into a large planning group which was charged with developing “provocative propositions”. Provocative propositions were statements of organisational aspiration and intent, based on the analysis of the organisation at its very best. These propositions were then validated by organisational members along two dimensions: (1)

⁵⁸ Options (2011) ‘Participatory Ethnographic Evaluation and Research’ from www.options.co.uk/peer.

⁵⁹ Johnson, P. C. and D. L. Cooperrider (1991) ‘Finding a path with heart: Global social change organizations and their challenge for the field of organization development’, *Research in Organizational Change and Development*, 5: 233-284.

how much does this statement capture our values? and (2) how much are we like this?⁶⁰.

g) Horizontal evaluation

This is another participatory approach that fosters programme improvement and knowledge sharing among network participants; combining self-assessment with external reviews by peers. It is a good method for organisations that work in different sites. By introducing peers it ‘neutralises the lopsided power relations that prevail in traditional external evaluations creating, a more favourable atmosphere for learning and improvement’⁶¹.

Horizontal evaluation has been used by farmers to evaluate ARD methodologies. It consists of the following:

The central element of a horizontal evaluation is a workshop that brings together a group of local participants, who are developing a new research and development (R&D) methodology and a group of peers from other organisations and mostly from other countries in the region, who are interested in learning from the experience. The workshop combines presentations about the methodology, field visits, small group work and plenary discussions. Views of the two groups concerning the strengths and weaknesses of the methodology under development are elicited and confronted. Suggestions are generated for improving the methodology, and further development of the methodology is encouraged in other settings.

h) Social Return on Investment (SROI)

Another way of assessing impact is the Social Return on Investment (SROI). This is currently used (among other methods) by African Women in Agricultural Science and Development⁶². SROI is an analytical tool to measure and account for a broad concept of value. It incorporates social, economic and environmental costs and benefits into decision-making, providing a fuller picture on how value is created or destroyed. SROI is a measuring methodology built on traditional

⁶⁰ Bushe, G. R. (1995). ‘Advances in Appreciative Inquiry as an Organization Development Intervention,’ Organization Development Journal 13(3): 14-22.

⁶¹ Thiele, G., Devaux, A., et al. (2006). ‘Horizontal evaluation: stimulating social learning among peers’, Institutional Learning and Change Initiative (ILAC) Brief 13:, p1.

⁶² <http://awardfellowships.org/home.html>

forms of cost-benefit analysis and aims to capture the economic value of social benefits. It transforms social objectives into financial measurements. SROI is based on seven principles:

1. Involve stakeholders - Understand the way in which the organisation creates change through a dialogue with stakeholders.
2. Understand what changes - Acknowledge and articulate all the values, objectives and stakeholders of the organisation before agreeing which aspects of the organisation are to be included in the scope; and determine what must be included in the account in order that stakeholders can make reasonable decisions.
3. Value the things that matter - Use financial proxies for indicators in order to include the values of those excluded from markets in same terms as used in markets.
4. Only include what is material - Articulate clearly how activities create change and evaluate this through the evidence gathered.
5. Do not over-claim - Make comparisons of performance and impact using appropriate benchmarks, targets and external standards.
6. Be transparent - Demonstrate the basis on which the findings may be considered accurate and honest; and show that they will be reported to and discussed with stakeholders.
7. Verify the result - Ensure appropriate independent verification of the account.

Recent donor pressures for measuring results may be a key push for this type of approach that aims to include social change in financial measures. The quality of the exercise will depend not on churning out financial measures, but on the richness and inclusiveness of the debate to identify, measure and interpret the proxy indicators for this ‘financialisation’.

There are many other participatory approaches based on a similar logic to the tools detailed above: self-assessment over external ‘arms length’ approaches, qualitative data directed at learning for organisational change (and thus delivering practical useful outcomes) and empowering participants e.g. utilisation-focused evaluation and guided self-assessment (from the roots up).