

This article was downloaded by: [115.85.25.194]

On: 31 March 2015, At: 20:30

Publisher: Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office:
Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Gender & Development

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/cgde20>

A review of approaches and methods to measure economic empowerment of women and girls

Paola Perezniето & Georgia Taylor

Published online: 26 Jun 2014.



CrossMark

[Click for updates](#)

To cite this article: Paola Perezniето & Georgia Taylor (2014) A review of approaches and methods to measure economic empowerment of women and girls, *Gender & Development*, 22:2, 233-251, DOI: [10.1080/13552074.2014.920976](https://doi.org/10.1080/13552074.2014.920976)

To link to this article: <http://dx.doi.org/10.1080/13552074.2014.920976>

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at <http://www.tandfonline.com/page/terms-and-conditions>

A review of approaches and methods to measure economic empowerment of women and girls

Paola Perezniето and Georgia Taylor

This article presents findings from a review of 70 evaluations of development interventions which had direct or indirect impacts on the economic empowerment of women and girls. We defined this as a process whereby women and girls experience transformation in power, agency, and economic advancement. The review distilled knowledge about monitoring, evaluation, and learning methods and approaches being used. It recommends the use of mixed (quantitative and qualitative) methods to assess economic empowerment comprehensively. It also explored the development and use of relevant indicators to measure economic empowerment of women and girls; the rigour with which methods are used; patterns in the use of different approaches for distinct thematic areas (micro-finance, social protection, fair trade, legal frameworks, and so on); and approaches used to bring out the voices of women, men, girls, and boys, among other key questions. It considered strengths, weaknesses, innovations, and challenges involved in evaluation and research approaches and methods based on primary data. The review, and this article, aim to inform agencies commissioning evaluations on how to ensure women's economic empowerment dimensions are captured; and to help those designing interventions to ensure these support positive transformation in the lives of women and girls.

El presente artículo expone las conclusiones surgidas de una revisión de setenta evaluaciones aplicadas a varias acciones de desarrollo que tuvieron impactos directos o indirectos sobre el empoderamiento económico de mujeres y niñas. Las autoras lo definen como un proceso en el que las mujeres y las niñas experimentan una transformación en relación al poder, a su participación [agency] y a su avance económico. Dicha revisión produjo conocimientos acerca de los métodos de [monitoreo, evaluación, aprendizaje] y de los enfoques utilizados. En este sentido, recomienda el uso de métodos mixtos (cuantitativos y cualitativos) que permitan evaluar el empoderamiento económico. Asimismo, examina el desarrollo y el uso de indicadores que resulten adecuados para medir el empoderamiento económico de mujeres y niñas; el rigor con que fueron utilizados los métodos; los patrones de uso de distintos encuadres para las diferentes áreas temáticas (microfinanzas, protección social, comercio justo, marcos legales, etc.); y los enfoques utilizados para que las mujeres, los hombres, las niñas y los niños, pudieran expresar sus opiniones, entre otros asuntos clave. Apoyándose en datos primarios, la revisión consideró las fortalezas, las debilidades,

las innovaciones y los retos que conllevan los distintos enfoques y métodos de evaluación e investigación utilizados. Tanto la revisión como el presente artículo plantean la posibilidad de brindar información a las agencias que contraten servicios de evaluación respecto a cómo asegurar que se consideren las dimensiones del empoderamiento económico de las mujeres. Además, pretenden operar como auxiliares a la hora de implementar acciones, a fin de garantizar que éstas se dirijan a apoyar la realización de una transformación positiva en las vidas de las mujeres y las niñas.

Cet article présente les constatations d'un passage en revue de 70 évaluations d'interventions de développement qui ont eu des impacts directs ou indirects sur l'autonomisation économique des femmes et des filles. Nous avons défini cela comme un processus par lequel les femmes et les filles bénéficient d'une transformation en matière de pouvoir, de libre arbitre et de progrès économique. Le passage en revue a permis de condenser les connaissances relatives aux méthodes et approches de suivi, évaluation et apprentissage utilisées. Il recommande l'utilisation d'un mélange de méthodes (quantitatives et qualitatives) pour évaluer l'autonomisation économique de façon approfondie. Il a aussi examiné l'élaboration et l'utilisation d'indicateurs pertinents pour mesurer l'autonomisation économique des femmes et des filles ; la rigueur avec laquelle les méthodes sont utilisées ; les schémas de l'utilisation de différentes approches pour des domaines thématiques distincts (microfinancement, protection sociale, commerce équitable, cadres juridiques, etc.); et les approches utilisées pour faire entendre les voix des femmes, des hommes, des filles et des garçons, entre autres questions clés. Il a considéré les points forts, les points faibles, les innovations et les défis rencontrés dans les approches et méthodes d'évaluation et de recherche basées sur les données primaires. Le passage en revue, et cet article, ont pour but d'informer les organisations qui commanditent des évaluations sur la manière de veiller à bien rendre les dimensions de l'autonomisation économique des femmes; et d'aider ceux qui conçoivent les interventions à veiller à ce que ces dernières soutiennent une transformation positive dans la vie des femmes et des filles.

Keywords: evaluation; methodologies; gender; economic empowerment; women; girls; review

Introduction

The economic empowerment of women and girls is understood by us to be a process whereby women and girls experience transformation in power and agency, as well as economic advancement. This article is based on a review of evaluations of development programmes and projects that directly or indirectly have an impact on women's and girls' economic empowerment (WGEE). The objective of the review was to provide information to improve the design and commissioning of evaluations to deliver stronger positive impacts across the different dimensions of WGEE. Overall, the aim

was to inform agencies commissioning evaluations on how to ensure gender empowerment dimensions are captured; as well as those designing interventions on how to better foster the positive transformation of women and girls. The review was commissioned by the UK Government Department for International Development (DFID), and was carried out from March 2013 to March 2014.

The review was not based on primary data, but produced a meta-analysis based on many research and evaluation documents that did analyse primary data. The review explored strengths, weaknesses, innovations, challenges of evaluation and research approaches, and methods used in specific, grounded cases. The first stage identified 254 empirical evaluations and studies of mainstream and targeted interventions that measure or assess the effects of the intervention on WGEE. After a systematic screening process, a shortlist of 70 documents were analysed in detail by a panel of experts in order to distil knowledge about monitoring, evaluation, and learning methods and approaches being used to evaluate or analyse effects on WGEE.

The review explored a wide range of issues, including the use of quantitative and qualitative research methods in different types of development interventions with various aims. It also explored a range of key questions. These included the development and use of relevant indicators to measure economic empowerment of women and girls; the rigour with which methods are used; patterns in the use of different approaches for distinct thematic areas; and the development of legal frameworks around issues such as land and labour rights; and effective approaches that amplified the voices (that is, opinions and choices) of women and girls, whose voices are often absent in evaluation processes and reporting.

The review was wide in scope, analysing evaluations that included some measure of WGEE in one or more of the following thematic areas:

1. Financial services.
2. Business development services.
3. Skills training.
4. Asset provision (both financial and not financial).
5. Social protection.
6. Unions and fair employment.
7. Trade and access to markets.
8. Regulatory and legal frameworks.

This article shares some of the findings. We write it having developed the framework and methodology for the review, managed the review process, and undertaken the meta-analysis of expert reviews of the 70 shortlisted reports.

Defining and exploring WGEE

There is, of course, a wide literature focusing on women's empowerment as an economic, social and political process, and many writers have used a typology of power which identifies different forms of this, all of which are relevant in analyses of empowerment processes.¹ For the purpose of this review, we adapted aspects of this typology to shed light on WGEE specifically. We defined this as a process whereby women's and girls' lives are transformed from a situation where they have limited power and access to assets to a situation where they experience economic advancement, and their power and agency is enhanced (VeneKlasen and Miller 2002).

The four dimensions of power are referred to by us as 'change outcomes', and defined as follows:

1. *Power within*: the knowledge, individual capabilities, sense of entitlement, self-esteem, and self-belief to make changes in their lives, including learning skills to get a job or start an enterprise.
2. *Power to*: economic decision-making power within their household, community, and local economy (including markets), not just in areas that are traditionally regarded as women's realm, but extending to areas that are traditionally regarded as men's realm.
3. *Power over*: access to and control over financial, physical, and knowledge-based assets, including access to employment and income-generation activities.
4. *Power with*: the ability to organise with others to enhance economic activity and rights.

Empowerment is thus understood as a process of change that transforms women's and girls' lives in these four areas and interacts with resources (pre-conditions), agency (process), and achievements (outcomes) (Kabeer 1999). Naila Kabeer's view of empowerment places economic empowerment in the spotlight as an essential underpinning of wider social and political empowerment of women, both as individuals and as a collective marginalised group. Empowerment will only be possible and sustainable if there are changes at different levels: within the individual (capability, knowledge, and self-esteem); in communities and institutions (including norms and behaviour); in available resources and economic opportunities; and in the wider political and legal environment (Golla *et al.* 2011).

Trends and gaps from the initial search

Based on the 254 studies included in the database, we can identify the following trends and gaps.

Trends

- Just under half (46 per cent) of the evaluations and studies focus primarily on financial services and their effects on WGEE. Within this thematic area, interventions supporting micro-credit and self-help groups (including savings and loans schemes) are commonly evaluated, but other areas such as micro-insurance and the use of new technologies (such as mobile phones) are less well explored.
- Interventions (particularly successful ones) promoting WGEE generally combine services across different thematic areas.
- There was a relatively even balance in terms of the methods used, with 30 per cent of the evaluations using quantitative methods, 38 per cent qualitative methods, and 28 per cent using mixed methods. Less than one in ten (9 per cent) used a different type of method or methods that were unclear. Among quantitative evaluations, the most common approaches to be used were randomised control trials (RCTs) and studies using quasi-experimental design.
- In geographical terms, around one-third of the reports covered interventions in South Asia (35 per cent), with a further one-third (30 per cent) focusing on interventions in sub-Saharan Africa.

Gaps

- Only one-fifth (21 per cent, or 55) of the evaluations and studies included in the database explored economic empowerment outcomes for adolescent girls. While these kinds of interventions have been increasing in recent years with initiatives such as the Girl Hub,² there are still few documented evaluations of their impacts.
- Most of the reports lacked age-disaggregated data, therefore failing to recognise the different experiences of girls and women during different stages of the life cycle.
- In terms of thematic areas, there were relatively few reports in three areas: legal and regulatory frameworks (two reports); unions and fair employment (six); and asset provision (four). This might be because few interventions in these areas have targeted women, reflecting a gender-blind approach; or it could be because a dearth of gender-disaggregated data in these areas does not allow for a differentiated analysis of the effects of these interventions on women's and men's economic empowerment.

An overview of the review methodology

The methodology used drew on and adapted systematic review principles,³ in order to ensure a comprehensive search of the literature, and to identify as many relevant

resources as possible. Since the objective of the review was to explore evaluation methodologies and approaches, rather than findings (which are generally the focus of reviews), the tools used at every stage were adapted to capture the information needed to answer the research questions.

The review process had three stages: an initial literature search, narrowing the field of inquiry, and review.

Initial literature search

The search strategy was designed to identify evidence and evaluations from across disciplines that report outcomes for WGEE. The search was largely Web-based and designed to cover both grey and academic literature, without privileging any particular discipline. The systematic search process used consistent inclusion criteria, search terms, coding, and search locations. A total of 382 papers were uploaded into the database for review. Only the 254 that reported empirical research (excluding systematic reviews, literature reviews, and theoretical papers) were subsequently analysed for quality in the next stage.

Narrowing the field of inquiry

To narrow down the field from the 254 empirical reports identified in the first stage, the research team devised a two-stage process using two different tools. The first 'basic scoring tool' captured the extent to which there was information on the methodological approaches used in the report. Reports that scored two or more points on that tool (160 of them) were then assessed using a purpose-designed 'quality assurance tool'. This tool gave a rapid assessment of quality, based on basic criteria set out for each methodological approach used. It provided a light-touch analysis through a rapid review of the documents, enabling further in-depth screening. Reports that scored two or more points on this second tool were considered for the final stage of the review. In total, 70 documents 'passed' the quality assurance test.

Review

Out of the 70 documents shortlisted, there was at least one in each of the eight thematic areas. They spanned a range of methodological approaches. Thirty-one used quantitative methods only, and a further 26 used a mix of both quantitative and qualitative methods, with 13 using qualitative methods alone.

Shortlisted reports were reviewed by a panel of four experts, using a review template that contained questions derived from evaluation best practice, as well as key resources on women's economic empowerment – for example, the Development Assistance Committee (DAC) criteria for evaluating development assistance,⁴ Bond's

quality of evaluation checklist, DFID's evaluation policy (DFID 2013), and Kabeer's (1999) paper on reflections on the measurement of women's empowerment. The purpose of the review template was to enable the expert panel to make a robust analysis of the quality of methodologies used, particularly in relation to their ability to capture relevant information about the effects of the intervention on WGEE. The individual reviews and the synthesis provided by each expert form the basis for our analysis, findings, and recommendations.

What makes a good-quality evaluation that effectively measures impact on WGEE?

This section analyses key dimensions of 'quality' in the approaches used by the 70 shortlisted evaluations or study documents that were reviewed by the expert panel. It gives examples from specific evaluations to illustrate strengths or weaknesses in relation to the quality criteria defined during the review phase. It draws largely on information from the expert reviews, and builds on insights derived from the meta-analysis of the set of reviews.

Applying a holistic approach to WGEE

According to the definition of WGEE used for this review, it is necessary for the intervention to result in some kind of transformational change on *women and girls' economic advancement and women's and girls' power and agency*, in addition to increasing income: we term that a 'holistic' approach.

Not all of the 70 shortlisted documents reflected such an approach; many looked at change in terms of economic advancement only. In cases where change in women's power and agency was assessed, it was common to limit this to decision-making power within the household. The quantitative studies were less likely to have a multi-dimensional and holistic view of women's economic empowerment, possibly because the types of indicators used tend to be more limited to changes in economic status, and do not always capture how and why changes happened to women and/or girls.

Evaluations that demonstrated a holistic approach to WGEE included those authored by Sally Baden (2013), on Oxfam work, Supriya Garikipati (2008), Sirojuddin Arif *et al.* (2010), on Overseas Development Institute (ODI) work, and Nicola Jones and Mohammed Shaheen (2012), on ODI work. However, there are varying levels of sophistication of the WGEE frameworks used. At the very minimum, the evaluation would measure changes in the individual's decision-making power within the household as a measure of women's agency. *The lack of measurement of change in norms and attitudes* was generally not done at the institutional level and at the level of the legal and policy-enabling environment.

Building an evaluation team with gender expertise

The review found that evaluation teams with gender experts or expertise presented better quality analysis in relation to WGEE – a finding that perhaps is not so surprising to gender and development practitioners. In addition, gender expertise ensured integration of gender concerns into research methodology.

Preparing for the evaluation and understanding the context

Good-quality evaluations in the review always contained a good analysis of the social and economic context, including a strong gender analysis. If context analysis is done sufficiently well, it will enable a relevant theory of change and an evaluation methodology that takes into account the context. Good context analysis presented market and economic context, political and institutional context, gender analysis including data on the gender division of labour, roles and power relations, and information on wider processes of economic, social, and political change at household, community, and institutional levels.

Gender analysis is key to good context analysis. This is essential for good-quality evaluations, and our findings bore this out. Fewer than 25 per cent of the reports reviewed included a strong gender analysis. Gender differences, gender-related norms and behaviours, gender roles, and gender relations differ in every context, and assumptions should not be made when designing an evaluation. In some cases, an evaluation did aim to measure women's economic empowerment but there was a lack of understanding of the importance of the multiple gender inequality dimensions present, and how they related to the intervention's achievements, and its evaluation.

Most of the highest-scoring evaluations and studies had a well-articulated theoretical framework and/or theory of change. This enabled the evaluators to demonstrate where and how change would happen as a result of the intervention – and thereby to frame their evaluation around that change process. The best theories of change showed dynamic and multi-faceted change processes that reflected transformational changes in women's agency and economic advancement, and the relationship between them.

A useful theory of change can be seen in the evaluation of Henny Slegh *et al.* (2013) of a pilot project in Rwanda that targeted male partners of women involved in a CARE-funded microfinance scheme so as to counter gender norms, power dynamics, and gender-based violence. The report found that women participants in micro-credit programmes need to be supported by improved and better programming which engages with men in deliberate and structured ways, including promoting greater male involvement in care work and that in some settings, solely focusing on women may lead to negative effects for women, both in the short and long term. The theory of change employed situates the pilot project within broader literature on the multiple

spill-over effects of women's economic empowerment on their individual social well-being and on familial well-being. But they note that this positive relationship cannot be assumed and is, in part, dependent on men's reactions to women's changing roles and status.

Methodology and indicators

The review found that mixed-methods evaluations were the most effective in terms of measuring change and outcomes, and also provided good material for learning and improving future project design. This finding is in line with feminist literature on research methods (Oakley 2000). A good mixed-methods approach means that qualitative and quantitative data can complement each other, so that the whole is more than the sum of its parts. It also allows for effective triangulation⁵ of data. In low-resource cases and when mixed methods is not feasible, qualitative or quantitative data alone would be useful as long as the evaluation design was robust and includes 'holistic' indicators for gender economic empowerment, has a solid theory of change around transformational impacts of the evaluation on women and/or girls, and where a gender expert supports the research design and/or analysis of findings to ensure important aspects are analysed. There are complementary roles of both kinds of methods:

- *Quantitative* methodologies tended to be used to demonstrate that: (1) change had taken place; (2) the intervention caused the change to take place (*causality*); and (3) the findings can be generalised across a population group (in the case of representative samples).
- *Qualitative* methodologies were most commonly used for: (1) context analysis to design sampling and tools; (2) conducting participatory activities to identify indicators of change; (3) establishing how the change takes place; (4) understanding *why* change happens (or does not happen); (5) researching how people understand and describe that change; and (6) identifying unintended changes or impacts.

Sequencing of quantitative and qualitative data collection was relative to the particular role each method played in the overall evaluation. However, in the reports reviewed it was rare to see a good-quality explanation of why mixed methods had been chosen, and how a particular sequence had been decided.

Multiple sources of data were important for triangulation in cases where there was no baseline. Evaluations were generally scored higher if the analysis in a mixed-methods approach used both the quantitative and qualitative data to complement and reinforce each other. However, thorough triangulation in mixed-methods evaluations was rare;

many had separate and distinct sections for the qualitative and quantitative aspects, which in some cases did not reference each other.

Funding appeared to be lower for qualitative evaluations. A large number of those reviewed (or parts of them) had limited scope and reach, with small samples. This suggests a lack of value and funding attached to qualitative methods. In many of the studies reviewed, it is unlikely that the point of data saturation was reached.

For *measures of change*, the choice of dependent variables for quantitative analysis is vitally important for a good-quality assessment of change in WGEE.

In many cases, evaluations were measuring progress against a pre-agreed set of indicators that appeared to be at output level only, with no analysis of unexpected or wider changes that might have taken place. Measurement of change at outcome or impact level required a set of indicators that (given the context) adequately measured transformational change on women's and girls' economic advancement and power and agency.

The reviews identified that evaluation teams and researchers often make underlying assumptions about gender roles and relations, and women's activities, and use stereotypes. This can limit the level of analysis, for example leading to women and men being asked questions based on misconceptions:

- *Type of expenditure*: Women may only be asked about expenditure on household items or child-related expenditure, whereas men are asked about recreation and other activities. This reinforces the stereotype that women are supposed to be responsible for household and child-related expenditure, and may limit investigation into changes that could be taking place. It also negates the possibility that men may also have a responsibility for household expenditure.
- *Type of decisions*: Women are typically asked about who makes decisions about issues that are relevant to the division of household labour (that is, who decides how women's and children's time is allocated), but seldom asked about other critical areas of life such as decisions about who participates in community or civic events, how women's time is allocated to activities outside the household (labour or social), and the like. Only rarely may questions be asked about joint decision-making, with researchers tending to frame their questions as 'either/or' (e.g. are household decisions made by men or women?). In many cases, joint decision-making is the norm in many decision-making spheres. Even when this is revealed it is rare to see an analysis that explores the relative power balance between men and women who are making decisions together. In general, this level of detail was missing from the evaluations
- *Type of economic activities*: Assumptions may be made about women being home-based, or economically active only in certain sectors, so that their involvement in other sectors is not considered. Considerations must also be made of unpaid care

work or subsistence work which changes women's workload dramatically and completely affects their engagement or non-engagement with paid work.

Evaluations that scored well used a multi-dimensional range of indicators to measure women's and girls' economic advancement and changes in their power and agency. The most innovative studies used unique power variables that captured concepts of economic empowerment that one might not normally consider, such as whether a young girl was less likely to have unwanted sex. An example is a study by Oriana Bandiera *et al.* (2012), focusing on the Empowerment and Livelihood for Adolescents (ELA) programme – implemented by the non-government organisation BRAC Uganda – which is designed to improve the cognitive and non-cognitive skills of adolescent girls. The study aims to look at how this intervention uses a dual-pronged approach, aiming to affect adolescent girls' economic prospects, sexual behaviour, and life attitudes. The dual-pronged approach aims to (1) build life skills to build knowledge and reduce risky forms of behaviour, and (2) provide vocational training aiming to enable girls to establish small-scale enterprises.

The choice of what variables to use will obviously be context-specific; and ideally indicators will be defined in a participative way with and by women and men. As this is often not possible, the review highlights a number of innovative variables that could be considered for use. These draw on generic elements of an experience of lack of power – lack of control over resources being a foundational element – which echo the generic model of empowerment laid out by Naila Kabeer. The ways in which these generic elements are manifested in a particular context and for a particular woman will vary.

A good example of indicators used can be found in Ranjula Bali Swain and Wallentin's (2007) study of whether microfinance empowers self-help groups in India. This study used indicators that included: women's primary activity; access to independent savings; hypothetical response to possible verbal, physical, and emotional abuse; awareness of rights; and political activity. These indicators go beyond traditional 'decision-making' indicators and encompass other areas of women's control over their lives that are appropriate for the economic empowerment topic and context under evaluation.

An example of how indicators can be made relevant to particular women in a specific context through participatory planning methods is offered by the study of Syed Hashemi *et al.* (1996) on women's rural credit programmes in Bangladesh which used eight indicators (with sub-variables and weighting for each) that were developed through extensive observation, personal interviews with respondents in the ethnographic study villages and with credit programme staff, as well as from basic survey data. Responses were consolidated into the following eight indicators: mobility; economic security; ability to make small purchases; ability to make larger purchases; involvement in major decisions; relative freedom from domination by the family;

political and legal awareness; and participation in public protests and political campaigning. There was also a composite empowerment indicator by which a woman was classified as empowered if she had a positive score on five or more of the main eight indicators. In creating the empowerment indicators, the authors intentionally included a variety of specific actions or items in each one and made a minimal use of weights. All the operational measures of empowerment in the analysis reduce empowerment to dichotomous variables (empowered versus un-empowered). The cut-off points were based on percentage distributions for each dimension, with those classified as 'empowered' being around the 25th or 30th percentile for most dimensions.

Hardly any of the reviewed documents had included a list and description of the complete set of evaluation tools used, either in the main report or in annexes; in some cases, sets of questions were included in footnotes (an example is Oriana Bandiera *et al.* 2012, in a study of BRAC Uganda's ELA programme, discussed above). Some reviews with good-quality methodology had poor design and implementation of instruments, which led to poor-quality data and weak analysis. Data collection methods sections in the reports reviewed rarely discussed gender issues with respect to methodology and enumerators/interviewers.

Involvement of men, boys, and young women and girls in the evaluation

Many of the evaluations reviewed did not conduct interviews, surveys, or focus group discussions with men and/or young men, unless the project was specifically targeted at men as well. It could be argued that it is impossible to measure women's economic empowerment unless men's attitudes and behaviours are also taken into account. It is also true that issues including the relative contribution and therefore significance of women's income cannot be discerned – even though these issues are important to the question of empowerment within marriage and the family – without research into men's economic role within the household. This can be questioned in interviews with women but triangulation and direct involvement of men in interviews is likely to improve the reliability of data. The engagement of men was seen as critical in gaining acceptance for the strategies chosen to improve women's economic empowerment. Some of the evaluations had a general sample with men and women, but with no disaggregation of results, so women's empowerment was not easy to identify.

There were very few evaluations of interventions targeting girls' or young people's economic empowerment. Reports were excluded because they were either poor quality and/or lacked detail. In the other evaluations, lack of age-disaggregated data meant that impact on women and girls was not separated out, even though there are many young women and girls who are married and may form an important part of a sample. There was no evidence of differentiated approaches being used.

Inadequacies in the survey or interview methodology

Though women's voices were captured in most evaluations, data analysis and presentation of quantitative evaluations hardly ever reported women's voices verbatim. This is a shame since direct quotations not only enable a better sense of women's own perceptions and hence more accurate data, but also women's voices convey findings more powerfully.

The review showed that questions are sometimes based on recall, with insufficient triangulation to back up the findings. Factual information may not be remembered accurately. Most of the expert reviewers agreed that recall is subjective and can be unreliable, so other data sources are needed.

Recommendations

The main justification for diverting funds to evaluation from implementation of a WGEE intervention is to yield practical recommendations for increasing its contribution to women's economic empowerment and for learning within the sector. With this in mind, recommendations have been made here to improve evaluation quality and learning potential.

Overarching recommendations

Ideally, evaluation frameworks should be designed at the project design stage and should have a methodology and indicators that are relevant to the project's theory of change, and should include a baseline.

Clarity of purpose and a strategy for how the evaluation findings will be used are essential from the outset.

Different types of interventions may require different evaluation approaches but should be taken to avoid 'ticking the gender box' with an approach and analysis that is not sufficiently in-depth. It is important to ensure that projects are built around a full understanding of the gender context, even where a gender component appears to be an afterthought.

When designing mixed-methods evaluations, it is important to ensure that both the quantitative and qualitative components are well designed and able to complement each other, either by sequencing findings from one to inform the design of the other, or by ensuring that each approach collects complementary data.

The review identified very few evaluations and studies that looked at the economic empowerment of adolescent girls, despite a growing interest in this field. This implies a need for such interventions, but also for stronger evaluations and research in this area. Such interventions provide an opportunity to engage adolescent girls, and thus also present an opportunity to assess impact on this important age group.

Reports should include summary diagrams or tables. Excellent material and findings can be buried if they are not then translated into easily accessible formats.

It is useful if evaluations are part of a clear plan for dissemination of findings to deliver on the evaluation's purpose if this is to include learning and better implementation of the intervention in question (and similar interventions). Ideally, evaluation findings should be shared with research participants.

The report's recommendations should provide clear guidance for how to enhance impact on WGEE – and this should include policy implications. Policy recommendations could be enriched by drawing on evaluators' broader experience and knowledge in the field, not only the findings that come directly from the specific piece of research.

Build a team that has sufficient gender expertise and involve it from the beginning in evaluation design and context analysis

Evaluations should aim to include experts with solid experience and expertise in gender and economic empowerment/enterprise. Ideally, gender experts should know the technical area of the intervention being evaluated. If evaluators have expertise in a particular technical area but are not gender experts, the team's skills should be complemented with at least one gender expert, participating in the evaluation design stage but ideally also the data analysis stage.

Data collection with women, particularly on intra-household dimensions of economic empowerment, is best done by other women wherever possible in order to facilitate open discussions of constraints to women's agency. Where this is not possible due to the composition of the fieldwork team, the evaluation or study report should state that this was the case and note any possible response biases this might have caused.

Undertake context analysis that has a full gender analysis, and ensure there is a theory of change that describes transformational change in women's and girls' economic advancement and power and agency

Exploring women's economic empowerment (and the impact of any intervention) relies on a thorough context and problem analysis. The context analysis should cover gender equality and the economic, political, and technical context, so that the limitations of gender equality and its impact on women are fully understood.

Evaluation teams should not make assumptions about gender behaviour across contexts or allow underlying cultural and gender stereotypes held by the evaluation team, the data collection teams, or national partners to act as a barrier to good data collection and analysis, as this may ultimately limit the validity of the evaluation findings.

There should be a discussion of how political, economic, or natural situations and crises may have influenced the findings of the evaluation, particularly about WGEE.

It is essential to have a (explicit or implicit) theory of change that depicts the expected change processes, and guides the research questions and formation of indicators. This should include how transformational change is expected to take place in the short term, medium term and long term, and should cover changes in women's economic situation *and* agency.

Depending on the context analysis and the scope of the intervention being evaluated, the theory of change should show gender-related change expected at different levels: individuals, communities, institutions, and the legal and policy environment.

Use mixed methods effectively and ensure that indicators adequately measure transformational change

A set of minimum evaluation methodology standards would be a good way of ensuring quality. These could include some or all of the following.

Evaluations should provide a full explanation of the methodology and indicators used, including information about why certain choices were made and the limitations of the chosen methodology.

Mixed-methods evaluations should be used for WGEE interventions where possible. The evaluation should aim to establish whether change took place and what caused it, through quantitative analysis. However, qualitative information will be needed to understand the process of change, and how implementation has been experienced by individuals, as well as to uncover any unintended impacts or changes and other factors that may have influenced change. This is especially important for WGEE projects as changes in norms, attitudes, and behaviours are difficult to understand fully with quantitative data alone.

RCTs are generally considered the 'gold standard' of research methodology for evaluating impact, proving causality, and generalising findings. However, when resources or technical expertise are insufficient to do a comprehensive and statistically sound quantitative evaluation, other methods should be considered. Information can be more cheaply collected through use of good qualitative or participatory methods, together with a simple survey focusing on specific issues arising. Quasi-experimental design, drawing on propensity score matching (PSM) or difference in difference (DD) are a good alternative when RCTs are technically not feasible, because there can be no randomisation (e.g. when the evaluation is done after the intervention has been implemented), if robustly designed and implemented. They are best complemented by qualitative data.⁶

When using qualitative methods, it is important to include quotes from in-depth interviews and case studies as these add legitimacy to the findings and help to bring the issues alive for the reader.

Indicators and tools to be used in the evaluation should be based on local knowledge. Conduct prior and follow-up participatory and/or qualitative research to identify appropriate local indicators and ways of wording questions so that they are likely to be clearly understood and taken seriously by interviewees and thus encourage the most reliable responses.

Indicators should include measurement of the following:

- transformational change in women's economic advancement and agency;
- attitudes, norms, and behaviours of women and men;
- change in gender mainstreaming and attitudes in institutions;
- gender differences in empowerment between women and men to increase local relevance;
- changes in men's and women's social capital, economic autonomy, and political participation;
- rights of women in enterprise;
- effects of the wider market, value chain, and employment, as all of these will affect women's economic advancement and empowerment.

Ensure that sampling selections are explained and justified, and that data are disaggregated fully by age, sex, and other relevant groups

The sample should include women and men, and, where appropriate, girls and boys – even if the intervention is targeted at women and girls only. Evaluators and researchers should analyse the difference, the co-dependence between the groups, the positive outcomes, and relationships between men and women – not just the outcomes for women.

Evaluations should differentiate between women of different ages and their stages in the life cycle, as well as women from different backgrounds, even if the conclusion is that these different groups of women need something apart from the project activities; findings might yield important information about complementary services that would need to be put in place to maximise impact.

Appropriate tools should be designed for different study participants and provided in full in the research report

Evaluators should ensure that the evaluation/study tools are available in the final report (as an annex) for transparency and replicability.

The team should design appropriate tools and data collection methods for young people and different population groups in the sample. This is particularly important when collecting sensitive data from adolescent girls and boys. Participatory tools can work well with this cohort.

In order to demonstrate effect and especially impact in the absence of baseline data, there should be greater use of *historical tools* – e.g. life histories, generational comparisons, community histories.

Paola Perezniето is a Research Associate in the Social Development Programme, Overseas Development Institute. Postal address: 203 Blackfriars Road, London SE1 8NJ, UK, as well as a freelance consultant. Email: p.perezniето@odi.org.uk

Georgia Taylor is a freelance consultant and a director of WISE Development, Central Hall Westminster. Postal address: Storey's Gate, London SW1H 9NH, UK. Email: georgiat@btinternet.com

Notes

- 1 The four dimensions of power referred to here have been developed and adapted from earlier work which is well-referenced in the literature. For example, the concept of 'power-over' particular individuals or groups was first conceptualised by Stephen Lukes (1974) as structural domination – a quite different usage from the positive sense in which it is used in this typology used in the review. The other dimensions of power are used widely in writings on women, power, and development (e.g. Kabeer 1994; Rowlands 1997; Townsend *et al.* 1999).
- 2 A joint initiative between DFID and the Nike Foundation.
- 3 Although first applied in the medical sciences in the 1970s, systematic reviews have been recently, and increasingly, used in the field of international development to examine the impacts of a range of development and humanitarian interventions. Systematic reviews are a rigorous and transparent form of literature review. Described as 'the most reliable and comprehensive statement about what works', systematic reviews involve identifying, synthesising, and assessing all available evidence, quantitative and/or qualitative, in order to generate a robust, empirically derived answer to a focused research question (Mallett *et al.* 2012); also 'DAC Criteria for Evaluating Development Assistance', www.oecd.org/dac/evaluation/daccriteriaforevaluatingdevelopmentassistance.htm (last checked by the authors May 2014).
- 4 The four experts on the panel have significant experience of issues of gender equity and women's empowerment, as well as robust knowledge of quantitative and qualitative research methodologies. Nevertheless, when assessing the quality of a report, panel members were influenced by their individual expertise, knowledge base, and methodological preference. This review presents a balanced analysis based on careful reading of all the individual reviews and the analytical syntheses prepared by each expert.
- 5 Triangulation refers to the use of more than one approach to the investigation of a research question in order to enhance confidence in the ensuing findings. Since much social research is founded on the use of a single research method and as such may suffer from limitations associated with that method or from the specific application of it, triangulation offers the prospect of enhanced confidence.

- 6 Impact evaluations estimate programme effectiveness usually by comparing outcomes of those (individuals, communities, schools, etc.) who participated in the programme against those who did not participate. The key challenge in impact evaluation is finding a group of people who did not participate, but closely resemble the participants had those participants not received the programme. Measuring outcomes in this comparison group is as close as we can get to measuring 'how participants would have been otherwise'. There are many methods of creating a comparison group. Some methods are better than others. From a quantitative research perspective, all else equal, randomised evaluations (or randomised control trials) are the most effective. They generate a statistically identical comparison group, and therefore produce the most accurate (unbiased) results. There are 'quasi-experimental' methods that can be used when randomisation is not possible, and a different control group needs to be identified. Some of the common techniques used by this approach include PSM and DD. A detailed explanation of when different quasi-experimental methods should be used is found at: www.povertyactionlab.org/sites/default/files/documents/Experimental%20Methodology%20Table.pdf (last checked by the authors May 2014).

References

- Arif, Sirojuddin, Muhammad Syukri, Rebecca Holmes, and Vita Febriany** (2010) *Gendered Risks, Poverty, and Vulnerability: Case Study of the Raskin Food Subsidy Programme in Indonesia*, London: Overseas Development Institute
- Baden, Sally** (2013) *Women's Collective Action: Unlocking the Potential of Agricultural Markets*, Oxford: Oxfam GB for Oxfam International
- Bandiera, Oriana, Niklas Buehren, Robin Burgess, Markus Goldstein, Selim Gulesci, Imran Rasul, and Munshi Sulaiman** (2012) *Empowering Adolescent Girls: Evidence from a Randomized Control Trial in Uganda*, London School of Economics, <http://econ.lse.ac.uk/staff/rburgess/wp/ELA.pdf> (last checked by the authors May 2014)
- Department for International Development** (2013) 'International Development Evaluation Policy', May 2013, London: DfID
- Garikipati, Supriya** (2008) 'The impact of lending to women on household vulnerability and women's empowerment: evidence from India', *World Development* 36(12): 2620–42
- Golla, Anne Marie, Anju Malhotra, Priya Nanda, and Rekha Mehra** (2011) *Understanding and Measuring Women's Economic Empowerment. Definition, Framework and Indicators*, Washington, DC: International Centre for Research on Women
- Hashemi, Syed, Sydney Ruth Schuler, and Ann P. Riley** (1996) 'Rural credit programs and women's empowerment in Bangladesh', *World Development* 24(4): 635–53
- Jones, Nicola and Mohammed Shaheen** (2012) *Transforming Cash Transfers: Beneficiary and Community Perspectives of the Palestinian National Cash Transfer Programme*, London: Overseas Development Institute
- Kabeer, Naila** (1994) *Reversed Realities: Gender Hierarchies in Development Thought*, London and New York: Verso

- Kabeer, Naila** (1999) 'Resources, agency, achievements: reflections on the measurement of women's empowerment', *Development and Change* 30(3): 435–64
- Lukes, Stephen** (1974) *Power: A Radical View*, London: Macmillan
- Mallett, Richard, Jessica Hagen-Zanker, Rachel Slater, and Maren Duvendack** (2012) 'The benefits and challenges of using systematic reviews in international development research', *Journal of Development Effectiveness* 4(3): 445–55
- Oakley, Ann** (2000) *Experiments in Knowing: Gender and method in the social sciences*, Cambridge: Polity Press
- Rowlands, Jo** (1997) *Questioning Empowerment: Working with Women in Honduras*, Oxford: Oxfam GB
- Slegh, Henny, Gary Barker, Augustin Kimonyo, P. Ndolimana and Matt Bannerman** (2013) "'I can do women's work": reflections on engaging men as allies in women's economic empowerment in Rwanda', *Gender & Development* 21(1): 15–30
- Swain, Ranjula Bali and Fan Yang Wallentin** (2007) *Does Microfinance Empower Women? Evidence from Self Help Groups in India*, Working Paper No. 2007:24, Uppsala: Uppsala University, Department of Economics
- Townsend, Janet, Emma Zapata, Jo Rowlands, Pilar Alberti and Marta Mercado** (1999) *Women and Power: Fighting Patriarchies & Poverty*, London: Zed Books
- VeneKlasen, Lisa and Valerie Miller** (2002) *A New Weave of Power, People and Politics: The Action Guide for Advocacy and Citizen Participation*, Bourton on Dunsmore, Rugby: Practical Action Publishing