

Findings and Lessons from a Meta-review of Evidence on the Impacts of Interventions to Promote Entrepreneurship in Developing Country Contexts

Journal of Entrepreneurship and
Innovation in Emerging Economies

2(2) 1–24

© 2016 SAGE Publications

India Private Limited

SAGE Publications

sagepub.in/home.nav

DOI: 10.1177/2393957516647263

<http://eie.sagepub.com>



Rae Wolpe¹

Abstract

Strong evidence on the effectiveness of small business support interventions is crucial to enhance the development of small businesses and their economic development contributions. A meta-review of the evidence contained in four systematic reviews evaluating the effectiveness of over 100 interventions was conducted in order to identify the main lessons regarding what has worked well and what has not worked well. Opportunities to improve the way in which systematic reviews are carried out have been identified. With respect to the impacts of business development interventions, the systematic reviews find positive impacts on various levels of outcomes, but that it is difficult to generalise these findings for a number of reasons. There is a need to monitor business outcomes over a longer period of time, as well to include intervention cost effectiveness information to better inform policy makers. There is also a need for clearer framework or typology for business development interventions, as well as a clear framework to guide impact evaluations and for reporting on job creation/ destruction impacts. The use of theory-driven approaches to evaluating entrepreneurship and business development interventions can improve understanding of the often complex pathways or causal linkages which may ultimately lead to job creation and/or preservation. There are often lengthy time-frames of a few years between the implementation of interventions and the ways in which these may impact on entrepreneurs and businesses. Evaluation studies will need to factor these time-frames into their design and increasingly collect longitudinal data.

Keywords

Entrepreneurship, business development support evaluation, systematic review, meta-review

Introduction

Globally, small business development and entrepreneurship is recognised as a key driving force responsible for accelerating economic growth, job creation and poverty reduction (International Labour

¹ Development Economist and Managing Director, Impact Economix, Mowbray, South Africa.

Corresponding author:

Rae Wolpe, Impact Economix, PO Box 13446, Mowbray 7705, South Africa.

E-mail: rae@impacteconomix.com

Organization, 2010; World Bank, 2013). Obtaining rigorous evidence on what is working well, and what is not, to effectively promote entrepreneurship and business development has become increasingly critical to governments and funders as well as to development practitioners. The field of evaluation holds much promise in providing rigorous evidence on the impacts of initiatives aimed at promoting entrepreneurship and business development. A meta-review of such evidence can contribute towards improved knowledge in this area.

A systematic review is 'a review of the research literature using systematic and explicit accountable methods' (Gough, Oliver & Thomas, 2012a). A meta-review is a review 'of' or 'about' other reviews (sometimes also referred to as umbrella reviews or reviews of reviews) (Gough et al, 2012a; Thomson, Russell, Becker, Klassen, & Hartling, 2011). Systematic reviews differ from traditional literature reviews in that they employ a systematic methodology for finding relevant literature, for including or excluding certain literature, and for synthesising and reporting the results (Gough et al, 2012a). Systematic reviews bring together all existing research studies focused on a specific question or intervention as a shortcut to the literature. Specifically, a systematic review integrates and interprets the studies carried out relevant to the research question(s).

There is a growing body of international and local evidence on what is working (as well as what is not working well) to support entrepreneurship, and enterprise development and systematic reviews are analysing the evidence and capturing the learnings from a wide range of impact evaluations as well as other research literature (e.g., see Attanasio, Augsburg, de Haas, Fitzsimons, & Harmgart, 2012; Augsburg, de Haas, Harmgart, & Meghir, 2012; Banerjee, Duflo, Glennerster, and Kinnan, 2009; Cho & Honorati, 2013; Crépon, Devoto, Duflo, & Parienté, 2011; Grimm & Paffhausen, 2014; Karlan & Zinman, 2010, 2011; McKenzie & Woodruff, 2012; Zandniapour, Sebstad, & Snodgrass, 2004).

Research Objectives

The objective of this paper is to synthesise and summarise the empirical evidence on the effectiveness of five types of small business support interventions (access to finance, entrepreneurship training, business development services including wage subsidies and incentives, improvements to the business environment and formalisation of informal businesses) in terms of their potential to impact on a range of economic development outcomes (e.g., job creation or retention, or new venture creation).

The purpose of the research is to contribute towards efforts and policy discussions aimed at improving small business support and development in developing country contexts through a better understanding of what kind of business support interventions have worked, or not worked, well and in what contexts (and to a certain extent with what target groups).

The main research question that this meta-review will attempt to answer is 'What can systematic reviews of small business support interventions tell us about what kinds of interventions work, and do not work (well), in developing country contexts and for whom in terms of the nature of their impacts?'

Literature Review

The purpose of this literature review is to both define what systematic reviews and meta-reviews are, as well as to identify the main procedures to be followed in conducting systematic reviews and meta-reviews in order to inform the meta-review approach and methodology used in this paper.

What Is Evaluation and What Are Systematic Reviews?

Evaluation can be defined in various ways but includes assessing the worth or value of something in a rigorous and systematic manner so as to produce strong and reliable evidence to inform decision making and learning. In the context of development, the following OECD (OECD, 2002) definition of evaluation is a useful one:

The systematic and objective assessment of an ongoing or completed project, program or policy, its design, implementation and results. The aim is to determine the relevance and fulfillment of objectives, development efficiency, effectiveness, impact and sustainability.... Evaluation also refers to the process of determining the worth or significance of an activity, policy or program. An assessment, as systematic and objective as possible, of a planned, ongoing, or completed development intervention.

Critical appraisal and synthesis of research findings in a systematic manner first emerged in 1975 under the term ‘meta-analysis’ (Gough, Oliver & Thomas, 2013). Some meta-evaluations are for the most part a quality assurance check of the approaches followed by previous studies, while others are the aggregation of data from existing evaluations.

Various definitions of systematic reviews exist with examples including the following:

- A review of the research literature using systematic and explicit accountable methods (Gough et al., 2012a).
- Overview of primary studies that use explicit and reproducible methods (Greenhalgh, 1997).
- A synthesis that takes a systematic approach to searching, assessing, extracting and synthesising evidence from multiple studies (www.betterevaluation.org).

Systematic reviews differ from traditional literature reviews in that they employ a systematic methodology for finding relevant literature, for including or excluding certain literature, and for synthesising and reporting the results (Gough et al., 2012a). Systematic reviews synthesise all the existing high-quality evidence using transparent methods to give the best possible, generalisable statements about what is known. According to Gough, Oliver and Thomas (2013),

Systematic reviews enable us to establish not only what is known from research, but also what is not known. They can inform decisions about what further research might be best undertaken, thereby creating a virtuous cycle. They enable researchers, policymakers and practitioners to answer key questions: ‘what do we know, how do we know it?’ and ‘what more do we want to know and how can we know it?’

Systematic reviews examine the methods, assess the quality and collate the results from available primary research studies addressing a particular question and conforming to pre-specified eligibility criteria (Gough & Elbourne, 2002). The aim of such research synthesis is to develop a more comprehensive and unbiased summary of research than would be possible by consulting individual studies, traditional literature reviews or by eliciting expert opinion (Caird, Sutcliffe, Kwan, Dickson, & Thomas, 2015). A systematic review of effects or impacts may also involve meta-analysis or the statistical pooling of summative information on study effect sizes. In order to better understand differences in findings by context, theory-based systematic reviews will use an explicit theory of change, collecting data on outcomes along the causal chain.

Gough, Thomas & Oliver (2012b) have also developed the following examples of review types (Figure 1).

Predominant Review Type	Review Questions
<i>Aggregative</i>	
'What works?' reviews	What is the effect of a health or social intervention?
Diagnostic test	What is the accuracy of this diagnostic tool?
Cost benefit	How effective is the benefit of an intervention relative to its cost?
Prevalence	How extensive is this condition?
<i>Configurative</i>	
Meta-ethnography [4]	What theories can be generated from the conceptual literature?
Critical interpretative synthesis [8]	What theories can be generated from the conceptual literature?
Mate narrative review [11]	How to understand the development of research on an issue within and across different research traditions?
<i>Configuring and aggregative</i>	
Realist synthesis [9]	What is the effect of a social policy in different policy areas?
Framework synthesis [25]	What are the attributes of an intervention or activity?

Figure 1. Examples of Review Types

Source: Gough et al. (2012b).

The current proliferation of different types of systematic reviews is viewed by some (Gough et al., 2012b) as creating challenges for the terminology for describing such reviews. This terminology is seen to be important for enabling the conduct and use of reviews as well as for further developing systematic review methodology. Gough et al. (2012b) propose the development of terminology for the main dimensions of variation in systematic reviews and identify the following three main dimensions of variation:

1. Aims and approaches (including what the review is aiming to achieve, the theoretical and ideological assumptions, and the use of theory and logics of aggregation and configuration in synthesis).
2. Structure and components (including the number and type of mapping and synthesis components and how they relate).
3. Breadth and depth and the extent of 'work done' in addressing a research issue (including the breadth of review questions, the detail with which they are addressed, and the amount the review progresses a research agenda).

What Are Meta-Reviews?

A meta-review is a review 'of' or 'about' other reviews (sometimes also referred to as umbrella reviews, reviews of reviews or overviews of reviews) (Gough et al., 2012b; Thomson et al., 2011). A 'review of reviews' is defined as 'a systematic review where the data are from other reviews' (Gough et al., 2012, p. 260). I will use the term meta-review for the sake of consistency throughout this paper to refer to a review of systematic reviews.

The nature of this paper's meta-review is configurative in that it aims to synthesise the findings of selected existing systematic reviews to answer the review questions in order to provide a meaningful picture of what the research is telling us. In 2011, Thomson et al. (2011) noted that such overviews of reviews were 'a new and evolving publication type'.

Meta-reviews differ from systematic reviews in a number of ways, including the following:

1. Meta-reviews have a greater focus on breadth rather than depth of coverage (Lavis, 2009);
2. Meta-reviews should assess the quality of evidence contained in systematic reviews and critically assess if the included quality of evidence judgements have been made consistently across systematic reviews (Thomson et al., 2011); and
3. Meta-reviews can contain additional analysis of findings for comparison across systematic reviews (Thomson et al., 2011).

Caird et al. (2015) identify three factors when meta-reviews can be considered as an appropriate solution:

1. The review addresses a broad research question;
2. Results are required within a short timescale (e.g., less than 12 months); and
3. Resources do not allow the employment of enough manpower to expedite a full systematic review in the time available.

The term statistical meta-analysis is defined by Gough et al. (2012a) to mean 'the use of statistical techniques to aggregate the results of included studies' and is often used as a method of synthesis of effect sizes and power (the likelihood that an analysis will be able to detect a statistically significant event) from experimental trials in reviews of the effectiveness of interventions.

Overview of Systematic Review and Meta-review Procedures and Methodologies

This section describes the main process steps involved in systematic reviews and in meta-reviews in order to inform the meta-review approach adopted in this paper.

The key process steps involved in systematic reviews and meta-reviews are broadly similar and are as follows (Caird et al., 2014; Gough et al., 2012a; Petticrew & Robert, 2006; Thomson et al., 2011):

1. Identifying the research purpose, objective and question.
2. An explicit search strategy to identify all possible relevant evaluations and which is documented, including defining and applying clear inclusion/exclusion criteria for which systematic reviews to include or exclude.
3. Systematic coding, description and analysis of included studies to match or build a conceptual framework.
4. Quality and relevance assessment of the methodological quality of each systematic review and consideration of various sources of bias including publication bias.
5. Synthesis and reporting including presentation of findings.

This process is summarised in Figure 2.



Figure 2. Systematic Review Steps

Source: Gough et al. (2012a).

In conclusion, there are clear steps to be followed when conducting a systematic review. These include the need for a thorough literature search process which is documented in detail to allow for replicability and updating. In addition, careful research quality reviews are needed of evaluations to ensure that the evidence is strong enough to support robust conclusions.

Methodology

Research Approach and Design

This research uses a mix of methods in order to answer seven research questions (see Figure 3). Firstly, a literature review is conducted on key literature dealing with systematic review and meta-review methodologies. The purpose of this literature review is to inform a framework which is used to conduct

a meta-review of four business development systematic reviews. Secondly, four business development systematic reviews were then identified using the following inclusion and exclusion criteria:

- i. Systematic reviews should not include evaluations which are older than 10 years.
- ii. The evaluated interventions should be implemented either in low or middle income countries.
- iii. The evaluated interventions should ideally use a randomised controlled trial (RCT) or quasi-experimental design. As Waddington et al. (2014) note, 'Due to the logic of confounding, the use of a control or comparison group which receives no (or a different) intervention is usually a key to dealing with the attribution of an effect to the programme, though is not a sufficient condition due to selection bias. Nonetheless, and subject to appropriate risk of bias assessment, the reviewer may want to consider the inclusion of pre-test/post-test (reflexive control) designs when assessing changes in outputs (or outcomes immediately resulting from the intervention) along the causal chain' (2014, p. 367).
- iv. The evaluated interventions should be limited to the five types of business support intervention (access to finance, entrepreneurship training, business development services and research and development, wage subsidies, improvements to the business environment).

Since most business development systematic reviews have been undertaken by organisations involved in funding and/or implementing business development interventions (i.e., governments, donors and specialised agencies), a snow-ball technique was used starting with the most recently identified systematic review at that time (Grimm & Paffhausen, 2014). Applying the inclusion and exclusion criteria resulted in the following four systematic reviews being selected:

- a) Cho and Honorati (2013): Meta-analysis of entrepreneurship and value chain initiatives.
- b) Grimm and Paffhausen (2014): Interventions for employment creation, small medium systematic review.
- c) Mckenzie and Woodruff (2012): What are we learning from business training, an entrepreneurship evaluations around the developing world?
- d) Zandniapour et al. (2004): Review of evaluations selected enterprise development projects.

These four systematic reviews either used different definitions of developing countries or if a definition was not used included evaluations from certain specific continents as shown in Table 1.

This meta-review includes two main components: a descriptive overview of the broad parameters of the four systematic reviews, and a qualitative thematic analysis of the main findings from each systematic review with respect to the impacts of five types of business development interventions.

It must be noted that three of the four business development systematic reviews adopt a quantitative meta-analysis approach for assessing the included evaluations, and their findings are therefore based in part on an assessment of effect sizes and other research quality criteria applied to the included evaluations. One of the four systematic reviews combines the meta-analysis with theories of change for each of the five business development intervention to deepen the interpretation of the findings regarding the effectiveness of these five business development interventions.

The broad research approach followed is qualitative as the focus is on the qualitative analysis of both quantitative and qualitative data in order to summarise and compare findings from selected systematic reviews. The selected systematic reviews themselves contain a range of quantitative and statistical analysis and the main findings from this analysis will be summarised and compared, but without conducting further quantitative statistical analysis.

Table 1. Four Business Development Systematic Reviews: Development Context of Included Evaluations

Cho and Honorati (2013): Meta-analysis of entrepreneurship and value chain initiatives	Grimm & Paffhausen (2014): Interventions for employment creation small medium systematic review	Mckenzie & Woodruff (2012): What are we learning from business training, an entrepreneurship evaluations around the developing world?	Zandniapour. et al. (2004): Review of evaluations selected enterprise development projects
Developing countries: 1993–2013 (no definition)	Low and middle income countries as per World Bank definition from 1990 to 2013	This is not specified/defined in terms of geography or time-frames; however, appears to focus on evaluations conducted 2011–2012	Focuses on evaluations between 1995 and 2003 ‘in various geographic areas’, Africa, Asia, Middle East and North Africa, Latin America and the Caribbean and Transitional Countries

Source: Wolpe (2015).

Methods Literature Review: Systematic Reviews and Meta Reviews

ATLAS.ti² (version 7), a computer-aided qualitative data analysis software (CAQDAS) package, was used as a data management and analysis tool for the literature review, including the analysis of the four selected systematic reviews.

The literature on systematic reviews was obtained firstly by identifying the main books that have been written on this topic, as well as the main authors who are referenced in these books. Further material written by these authors was identified in key relevant journals which include the *Journal of Development Effectiveness, Evidence and Policy* and *Research Synthesis Methods*.

Given that very little has been written on the topic of meta-reviews (otherwise known as 'systematic reviews of systematic reviews'), the author contacted one of the leading authors who has written on the topic of systematic reviews of systematic reviews via email (Professor David Gough at the Evidence for Policy and Practice Information and Coordinating Centre (EPPI-Centre), located at the Social Science Research Unit at the Institute of Education, University of London) and requested his assistance in identifying any recent literature on this topic. This resulted in one additional article being identified on this topic (see Caird et al., 2015). This article has been written for the journal *Evidence and Policy*.

A combination of an inductive and deductive approach to coding the literature was followed. An analytic approach based on inductive coding moves from the specific to the general, so that particular instances are observed and then combined into a category or general statement (Wildschut, 2014 citing Elo & Kynga, 2008). Researchers immerse themselves in the data to allow new insights to emerge (Wildschut citing Hsieh & Shannon, 2005, p. 1279).

Method: Meta-review of Four Business Development Systematic Reviews

Business development systematic reviews were identified through a snowball technique. This began with reviewing the references of a recently published systematic review of business development systematic reviews (Grimm & Paffhausen, 2014) to identify other possible relevant systematic reviews as well as key authors. An internet search (Google as well as Google Scholar) was conducted using key terms and certain combinations of these (e.g., systematic review, meta-review, business support, business development). A few key peer-reviewed journals focusing either on small business and/or evaluation, as well as online databases, were also searched using these key terms. Given the time constraints on this research, the research was limited to four business development systematic reviews and no attempt was made to identify all possible systematic reviews.

The codes for the key systematic review steps were then applied to the four business development systematic reviews in order to describe the key elements of each systematic review. In addition, codes for the impacts of the given business development interventions as well as codes for key conclusions and further research were applied to each systematic review. This allowed for the key findings of each systematic review to be summarised regarding the findings on the effectiveness of the five business development interventions.

The research design, methodology and questions are summarised in Figure 3.

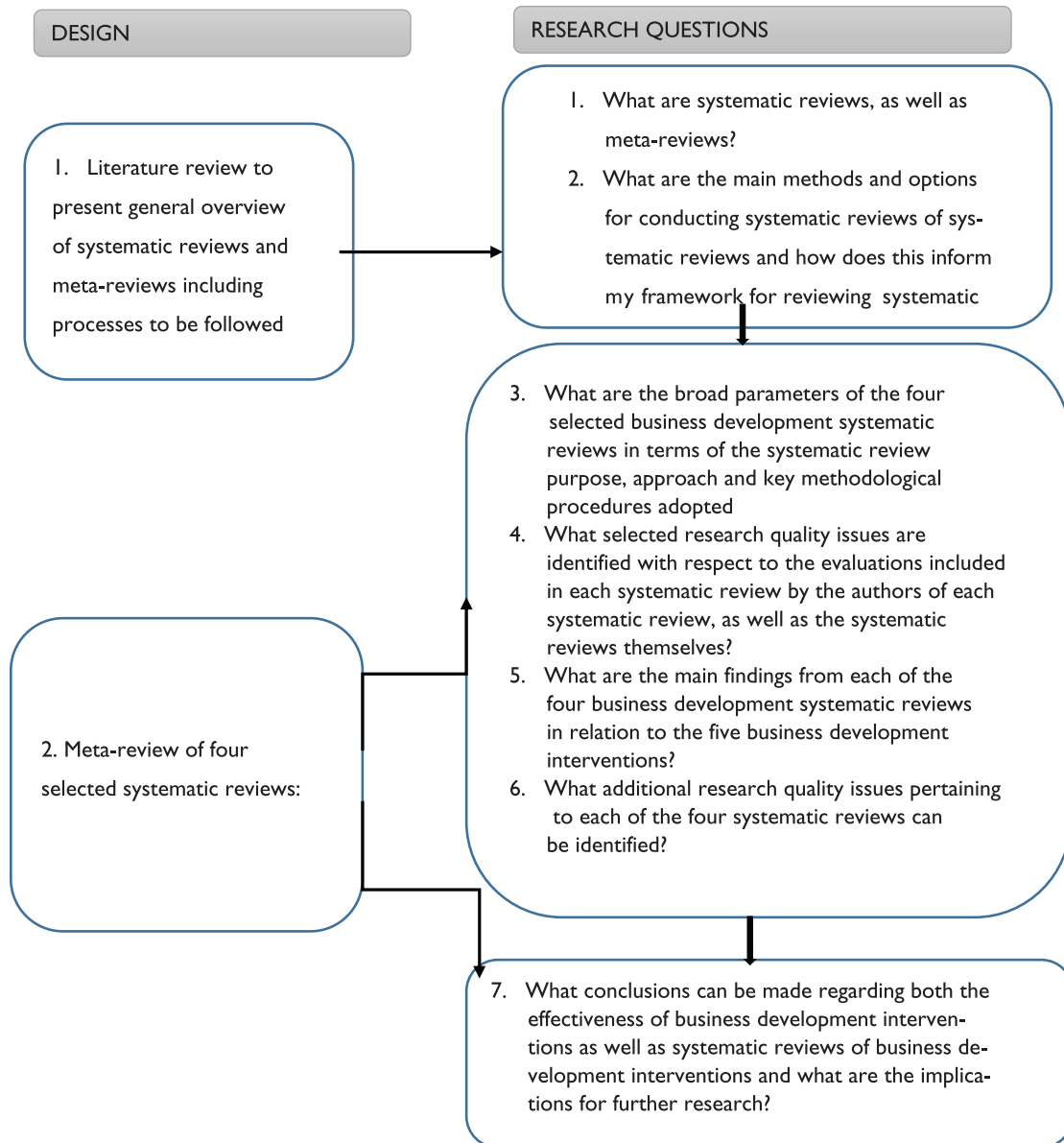


Figure 3. Research Design and Key Research Questions

Source: Wolpe (2014).

I have specifically included ‘nature of their impacts’ as part of the research question and not limited this to an examination of one kind of impact or outcome (e.g., employment) as it is recognised that business development interventions often have a variety of objectives, which may or may not include employment, and that there are numerous immediate and intermediate outcomes which are interrelated with employment impacts.

Findings/Results

Given that this meta-review is focusing on synthesising findings from systematic reviews which have examined the impacts/effects of business development interventions, the following inclusion and exclusion criteria have been used to guide the selection of systematic reviews:

- i. Systematic reviews should not include evaluations that are older than 10 years.
- ii. The evaluated interventions should be implemented either in low or middle income countries.
- iii. The evaluated interventions should ideally use an RCT or quasi-experimental design. As Waddington et al. (2014), note, 'Due to the logic of confounding, the use of a control or comparison group which receives no (or a different) intervention is usually key to dealing with the attribution of an effect to the programme, although it is not a sufficient condition due to selection bias. Nonetheless, and subject to appropriate risk of bias assessment, the reviewer may want to consider the inclusion of pre-test/post-test (reflexive control) designs when assessing changes in outputs (or outcomes immediately resulting from the intervention) along the causal chain'.
- iv. The evaluated interventions should be limited to the five types of business support intervention (access to finance, entrepreneurship training, business development services and research and development, wage subsidies, improvements to the business environment).

Meta-reviews can focus on a range of questions that are relevant to both evaluation and research methodology as well as to substantive evaluation findings. Examples of possible meta-review questions include the following (informed by Caird et al., 2015; Thomson et al., 2011):

- i. What are the broad parameters of the four selected business development systematic reviews in terms of the systematic review purpose, approach and key methodological procedures adopted?
- ii. What selected research quality issues are identified with respect to the evaluations included in each systematic review by the authors of each systematic review, as well as the systematic reviews themselves?
- iii. What are the main findings from each of the four business development systematic reviews in relation to the five business development interventions?
- iv. What additional research quality issues pertaining to each of the four systematic reviews can be identified (this question includes assessing issues related to each systematic review's main findings and therefore needs to be answered after consideration of those findings)?

For the sake of brevity, this paper will focus on findings relevant to Question iii above. Before summarising these findings, a brief overview of the four systematic reviews is presented.

Broad Parameters of the Four Business Development Systematic Reviews

Each of the four systematic reviews is described in the following terms with a brief commentary on key issues:

- i. Aim/purpose.
- ii. Inclusion/exclusion criteria for studies to assess.
- iii. Evaluation designs included/ excluded.

- iv. Population/target groups included/excluded.
- v. Types of business interventions included/excluded.
- vi. Outcomes of interest included.
- vii. Development context, time period and number of evaluations included.
- viii. Search strategies used and documented.

Aim/Purpose of the Four Systematic Reviews

The purposes of the four systematic reviews range from providing a theory-based approach to the question of which interventions create jobs and thus to contribute to the knowledge of how to create jobs, to identifying lessons from the evaluations as well as lessons for policy and future research, to improving our understanding of impact evaluation approaches being used for business development interventions.

Inclusion/Exclusion Criteria Used by the Four Systematic Reviews

There are wide-ranging differences in the various inclusion/ exclusion criteria used by each systematic review for what kinds of evaluations form the focus of each systematic review. With respect to the types of business development interventions included, there are many different types of interventions included. This reflects the fact that there does not seem to be one coherent framework which classifies the wide range of different types of business development interventions. There is also no standardised framework to classify the types of target groups or businesses linked to these different types of interventions (e.g., different definitions are used for micro, small and medium businesses in different countries in terms of different employee sizes or using other criteria such as turnover).

Evaluation Designs Included/Excluded

All of the systematic reviews limited themselves to experimental or quasi-experimental evaluations except for one (Zandniapour et al., 2004). This was to be expected given that three of the systematic reviews conducted meta-review on the impact results of the included evaluations. Only one also required panel or cross-sectional data at a firm level in order to be able to analyse changes in outcomes at different points in time (although it was noted that very few of the included evaluations had such data for any lengthy period of time after completion of interventions) (Grimm & Paffhausen, 2014).

Population/Target Groups Included/Excluded

Two systematic reviews identified entrepreneurs as one of the target groups included. However, there was no discussion of definitions of entrepreneurs. This is a vitally important issue as the literature identifies a number of different types of entrepreneurs and these different types have very different needs as well as very different potential levels of growth or future job creation potential. It is well known that many small businesses are not necessarily entrepreneurial.

The lack of consistent definitions for micro, small and medium businesses is another challenge noted by some of the systematic review authors. Only Grimm & Paffhausen (2014) review the range of existing definitions and applies a consistent firm size definition to the analysis. This lack of consistent definition application in both the included evaluations as well as across the systematic reviews makes it difficult to compare apples with apples, or compare interventions which focused on the same size range of businesses. This severely limits the generalisability and comparability of the findings related to specific types of interventions.

Types of Business Interventions Included/Excluded

There are widely differing types and descriptions of the types of interventions included in each systematic review. This is due to a number of reasons. Firstly, there is no standard classification framework of different types of business development interventions. Secondly, many interventions include multiple components, or cut across numerous intervention types (e.g., finance provision with training provision). The two most commonly included interventions related to the provision of various kinds of finance (e.g., micro-credit, grants, savings products, micro-insurance, guarantee funds, leasing products) and different kinds of training (e.g., business skills training, business plan development, financial literacy training and technical and vocational skills training). Other types of services included various business development services (e.g., consulting and counselling services, management and quality control practices, technology upgrading, market development services, export promotion services, information provision and networking services), research and development and product and process innovations, private sector incentive schemes, and enabling business environment changes in the policy, legal, institutional and regulatory conditions that govern business activities.

Outcomes of Interest Included

General categories of outcomes against which the impacts of business development interventions were evaluated relate to the following categories: labour market (both household impacts and individual impacts), business performance, business knowledge and practice, financial behaviours, attitudes (related to entrepreneurial traits). All the reviews included employment as one of the outcomes of interest. However, only one of the reviews (Grimm & Paffhausen, 2014) focused primarily on this outcome. Other outcomes included in one or more of the four systematic reviews were business start-up and survivorship, and business practices.

Importantly, there was a general lack of definitions for the outcome of employment as well as other outcomes. This issue is of fundamental importance as it is not possible to compare apples with apples unless apples have been clearly defined. For example, with respect to jobs, there are many different types of jobs (e.g., permanent, temporary, decent, high skilled, low skilled, high paid, low paid etc.). With respect to business profits, there are many different ways of calculating this (to be fair some of the systematic reviews do go into detail as to how profits were calculated in some of the individual evaluations reviewed). Thomson et al. (2011) highlight some of the methodological challenges regarding systematic reviews and analysing outcomes: 'For example, difficulties can arise when different sets of review authors report the same outcome in different ways, such as various time points or change from baseline versus end point data, or when the sets of outcomes in different reviews have minimal overlap.'

Development Context, Time Period and Number of Evaluations Included

Only one of the systematic reviews defines its geographic focus. The nature of the development context within which included evaluations have taken place is arguably a key meta-review issue and the lack of rigour in this regard also arguably limits the value of their findings. The general lack of relevant contextual information on the contexts within which the evaluations included in the systematic reviews take place in limits the extent to which any discussion of potential transferability of findings can take place.

The evaluations included the systematic review covered a 13-year period from 1990 to 2013. It is beyond the scope of this paper to discuss what time period can be deemed relevant with respect to the usefulness or not of business development evaluations and there is probably no clear answer to this. There are at least two relevant issues. The first is that ‘even if the range of systematic review (SR) is comprehensive and up-to-date at the time the overview of reviews is produced, this situation may change quite rapidly’ (Thomson et al., 2011). The second is that the macro-context at any specific point in time is an important external factor that can influence the effectiveness of business development interventions. So the global recession from 2008 onwards may have made conditions far more difficult for businesses in general, and impacted in different ways on different kinds of business interventions and different outcomes. This issue is only briefly mentioned in the one of the systematic reviews.

An analysis was conducted regarding which evaluations were included in more than one systematic review in order to get a sense of ‘double-counting’ lying behind the findings of the systematic reviews. This revealed the following number of duplicated evaluations—all in all there were about 11 evaluations that were included in at least two of the four systematic reviews out of a total of 138 evaluations included across the four systematic reviews (less than 10 per cent of this total). This is taken to mean that there are sufficiently numerous evaluations included across the systematic reviews to potentially include heterogenous results which in turn will allow for a rich comparison of results across the four systematic reviews.

Search Strategies Used and Documented

Only one of the systematic reviews contains a detailed search log (Grimm & Paffhausen 2014) of the results yielded from the different sources consulted, as well as the number of included and excluded evaluations from each source consulted. Only two of the systematic reviews mentioned that key scholars working in the field were consulted in order to identify the latest information on recently completed studies (Grimm & Paffhausen, 2014; McKenzie & Woodruff, 2012). Grimm & Paffhausen mention literature snowballing—this systematic review included by far the largest number of evaluations across the systematic reviews at 54.

Research Quality and Evidence Assessment Included in the Systematic Reviews

Systematic review authors need to assess how strong the evidence is in support of each of the overview’s conclusions, and present these assessments to the reader. The key limitations identified by the author’s of the four systematic reviews in their systematic reviews include the following:

- Challenges of attribution such as small samples sizes and low statistical power, as well as self selection all limit the ability to generalise the findings.
- The lack of firm-level data over sufficiently lengthy time periods to report changes in outcomes at different and over longer time periods.
- The lack of sufficient data for different target groups limits the ability to analyse which target groups benefit or not.

- Meta analysis cannot completely resolve the technical issues that are embedded in the original studies. So it is unclear whether a programme's performance is driven by its design or implementation (or both) from an individual study. By synthesising multiple studies, we can infer which design and implementation features are more associated with positive impacts, on average. However, because available information differs across studies, it is possible that the meta-analysis may omit important determinants of programme performance (Cho and Maddalena, 2013).
- The methodology of a systematic review, because of its focus on rigorous evidence, must systematically ignore untargeted policies, such as financial sector development, large-scale infrastructure projects, trade policies and alike, which may be particularly beneficial for small and medium enterprises. Other approaches must be applied to find out how effective these interventions are in creating employment (Grimm & Paffhausen, 2014).

Comparative Analysis of Systematic Review Findings by Five Kinds of Business Development Interventions

The key findings from the four systematic reviews are first summarised according to the following:

1. Access to finance.
2. Entrepreneurship training.
3. Business development services including incentives and wage subsidies.
4. Business environment improvements and formalisation of informal businesses
5. Findings specific to target groups (youth and women).
6. Overall findings.

Access to Finance

The main findings are that access to finance interventions have not been found to have much impact on employment for existing small and micro businesses but larger employment impacts have been achieved with both the creation of new (mostly micro) enterprises and the expansion of already larger, well-established and profitable firms.

In addition, there is also a long results chain from provision of finance, to how businesses use that finance and the time needed to potentially impact on employment. Finally, finance interventions targeting women have been found to be less successful at creating employment and the possible interpretation is provided that this is because women face additional constraints to be overcome when running businesses (it is beyond the scope of this paper to examine the range of these additional constraints in detail but they could range from family responsibilities to other gender challenges). Other findings were that:

- The details of the loan contracts matter. Field, Pande, and Papp (2010) find that short repayment periods, which over the loan period translate into lower outstanding loans and shorter maturities, prevent poor entrepreneurs from investing since they fear not being able to repay on time.
- Cash and in-kind transfers combined with training seem to yield larger impacts than microcredit (Cho & Honorati, 2013).

Entrepreneurship Training

The beneficial impacts of training have been found to include higher investment, process or product innovations, and/or growth in sales and revenue. Very few studies report impacts on profit and/or employment levels. Even non-existent or negative employment effects can be good news about trainings' effectiveness as entrepreneurship training was found to help non-profitable firms either to become profitable or to close down. Likewise, training can prevent non-profitable business ideas from being started (Grimm & Paffhausen, 2014). Relatively few studies have looked at how training affects the selectivity of who starts up a business or of which businesses survive. According to McKenzie and Woodruff (2012), evidence suggests training may enable less analytically able and poorer individuals to start businesses, and may prop up the survivorship of less profitable businesses. Training therefore appears to have some success in generating short-run impacts on business start-up. However, this does not necessarily increase employment, since it may come from people switching from wage work.

The length of training provided is an important issue influencing its impact, with a U-shaped relationship being found, suggesting that either intensive, short training or substantially extended training is best—however, the optimal solution depends on the goals of such interventions. The most successful training is tailor-made to address specific knowledge gaps. It appears that entrepreneurship training needs to be substantial in order to be effective, where substantial means that the training runs over an entire year with at least one training session per week (Grimm & Paffhausen, 2014). There were mixed results as to whether providing a package of training together with financial support is more effective than just providing training.

Most of the studies take a single snapshot of the impact of the training at a relatively short interval after training has ended. Two studies that have traced trajectories of impacts suggest that effects can indeed vary a lot over time. In de Mel, McKenzie and Woodruff (2012), the impacts on business start-up fade with time, as control firms catch up. Bloom, Eifert, Mahajan, McKenzie and Roberts (2012) find that introducing management practices in larger firms shows immediate effects on quality, then slowly leads to changes in inventory levels, output and productivity, and it is only after several years of using these practices that impacts start to show in terms of employment generation (through new plants being opened) (Grimm & Paffhausen, 2014).

The evidence is mixed on whether trainings' return is higher for those with initially lower skills. The review suggests, however, that entrepreneurship training is more helpful for start-ups than for business expansion. At the same time, it was found that the provision of individualised consulting support provided to larger firms can improve the performance of those firms (McKenzie & Woodruff, 2012).

Findings for Business Development Services (Including Wage Subsidies and Incentives)

With respect to wage subsidies, a demand-driven design where employers receive wage subsidies, as opposed to employees receiving vouchers, appears to work best at creating sustainable employment. In some cases, a supply-driven approach where workers receive the subsidy in the form of a voucher was found to have a negative impact on the employment of youth beneficiaries and only a few beneficiaries were retained in their job once the subsidy came to an end (Grimm & Paffhausen, 2014). There are a range of possible reasons for this including that there may be a better fit between employers and employees when employers choose the workers they hire (as opposed to being forced to hire a worker).

Firm conclusions regarding the impacts of incentives are difficult due to participant self-selection of firms into such programmes (i.e., that have a higher chance of succeeding than the general population). Nevertheless, tax breaks and fiscal incentives that are conditional on process and/or product innovations

seem to be particularly effective (however, the sample of studies is quite small and this finding must be interpreted with caution).

With respect to the provision of other business development services, findings included the following (Zandniapour et al., 2004). Business services were found to impact positively on enterprise sales, revenues, net profits or client income in 10 out of 15 evaluations that included these variables. However, the extent to which sales or revenues increased varied substantially across these studies and across enterprises within these studies. The findings also seem to suggest that relatively larger firms gain more in terms of increase in sales, profits or income compared to smaller, more micro firms. On the other hand, smaller firms generally attribute more of the increase in their sales/revenues and profits to their participation in the programme compared to larger firms. Projects that focused on the development of business service markets in general were found to have provided services that benefit clients and meet their needs. Projects generally have helped to remove internal firm constraints and increased enterprise sales, revenues and profits. In some cases, they also have contributed to the development of the market for business services (increasing the demand for and supply of services) by building local consultant capacity and increasing knowledge of small business needs and requirements. What has not really been established through these studies is the question of sustainability of the services provided or outreach of business services to the poor. Many of the studies address the issue of sustainability, but they do not entirely resolve the question of whether project interventions have led to the provision of quality services on a sustainable basis.

Findings for Business Environment and Formalisation of Informal Businesses

The main findings are as follows (Grimm & Paffhausen, 2014). Programmes that ‘force’ firms to formalise are unlikely to produce any significant employment effects as for many formerly informal firms formality does not translate into extra profits but into additional cost. In general, it seems easier to formalise firms while they are set up than formalising firms that already exist. Programmes that offer cheaper and easier formalisation procedures are more likely to have success but only for a relatively small group of entrepreneurs and firms that show already a higher initial performance. For the typical informal firm, it is not the costs of registration but the expected benefits of formality that is pivotal for their decision to formalise. So, the best incentive governments can provide for formalisation is to offer useful public services in return and in addition to simplifying administrative procedures.

Findings for Specific Target Groups (Women and Youth)

Youth-targeted programmes have shown good successes. For nearly all outcomes, particularly labour market activities and business performance, youth is highly associated with programme success. This is largely driven by youth-targeted programmes that present strong impacts (Cho & Honorati, 2013). Women face additional constraints which need to be addressed if support programmes are going to be successful. For women, the impacts from microcredit interventions seem higher than training programmes. Women are generally more severely credit constrained, and this in turn can hamper their potential gains from skills training. Financing support indeed performs better for women throughout all outcomes. Women are not associated with any large and significant impacts other than the outcome of attitudes, indicating that entrepreneurship programmes seem useful for female empowerment but may not be sufficient to address various barriers faced by women (Cho & Honorati, 2013).

Major Conclusions Contained in the Systematic Reviews

Overall, the systematic reviews point to difficulties in identifying impacts on job creation. The importance of addressing the broader macro/policy environment has been identified as key to enhancing employment

prospects for businesses. Overall, entrepreneurship programmes have a positive and large impact for youth and on business knowledge and practice, but no immediate translation into business set-up and expansion or increased income (Cho & Honorati, 2013). Many conditions have to be met before interventions in favour of individual enterprises do not only improve business performance but also lead to additional jobs. The results chain between enhanced knowledge and skills, and business investment, profitability and growth is a long and complex one and long time periods are needed to track changes in outcomes over time to better understand the dynamics impacting on these results chains. Not all potential and actual entrepreneurs can make good use of support. Different types of interventions will be required to increase employment for different groups. Interventions also need to pay attention to the interaction between different binding constraints. For instance, just improving business skills without facilitating access to capital (and vice versa), might often not be enough to have an effect on investment and employment (Grimm & Paffhausen, 2014).

Discussion

The four systematic reviews had a range of different, although closely related, aims. The methodologies adopted were similar in three of the four most recent systematic reviews, although the degree of rigour followed in all stages of the systematic reviews differed, with Grimm & Paffhausen (2014) following and capturing a far more rigorous search process.

Generally, the meta-analysis methods are suitable to examining issues of impact. However, insufficient utilisation of theory of change approaches (in the underlying evaluations) limited the analysis and findings in terms of better understanding what works for who, where, when and why. The use of the realist evaluation theory of change approach (see Pawson & Tilley, Pawson, 1997, 2004, 2006, 2013) in future evaluations could contribute towards a more detailed understanding of why certain interventions do or do not contribute towards achieving certain results or outcomes. Realist evaluation ‘Makes explicit the underlying theories or assumptions about how an intervention is supposed to work and evaluates theories against the available evidence’ (Pawson, Greenhalgh, Harvey, & Walshe, 2005). A realist review focuses on understanding and unpacking the mechanisms by which an intervention works (or fails to work), thereby providing an explanation, as opposed to a judgement, about how it works. ‘Mechanisms’ are understood to be a combination of resources offered by the social programme under study and stakeholders’ reasoning in response. Mechanisms can also be understood as ‘underlying entities, processes, or structures which operate in particular contexts to generate outcomes of interest’ (Astbury & Leeuw, 2010) or the process of how subjects interpret and act upon the intervention. Mechanisms are usually hidden (not visible), are sensitive to variations in context, and are responsible for generating outcomes (intended or unintended, positive or negative). An explicit focus on underlying generative mechanisms might help to counter what appears to be a growing trend towards oversimplified versions of programme theory in the form of linear logic models (Rogers, 2007; Weiss, 1997).

Another key gap or weakness in the four systematic reviews is their lack of a conceptual framework for business development, entrepreneurship and/or job creation. Such a framework would situate the range of business development interventions selected within a broader context of development dynamics, challenges and opportunities which impact on business development and job creation. This is important for policy-makers who need to understand the broad range of intervention options that are available to address particular policy problems. As a result, these systematic reviews do not directly add much value to the issue of understanding different intervention or policy options. This flaw is likely to exist in most

of the actual evaluations included in the systematic reviews themselves. One example pertains to the concept of the entrepreneurship ecosystem. This concept identifies a wide range of issues that need to be addressed to strengthen entrepreneurship. None of the systematic reviews mention these in any explicit or systematic manner, and only one systematic review notes that there are not many evaluations of broader business environment issues which impacts on the prospects for business growth (e.g., regulatory constraints) and therefore these are not taken into account in the systematic reviews.

In addition, possible concerns with publication bias may exist due to failure of all but one systematic review to include literature in a language other than English. Gough (2012) also notes that empirical research has shown that statistically significant 'positive' research results are more likely to be published (publication bias). Publication bias is systematic if the chance of publication was random, with no consistent impact on results.

It would appear that there has been ongoing strengthening of systematic review methodologies over time when comparing the four systematic reviews conducted between 2004 and 2014. Key systematic review process areas where improvements have been identified relate to the following:

1. Statement of systematic review purpose/aim/objective as well as the systematic review research question: This ranged from not being explicitly stated or made clear to explicitly stating the aim and research question.
2. Inclusion criteria: These have evolved from being quite broad to being very detailed and specific using the PICO framework, and to making fine-scaled differentiations that allow for the use of only selected evaluation results (and not necessarily all results) that fit the criteria.
3. Search strategies: These have also evolved from including just broad descriptions of sources consulted to including detailed search logs of all search results along the search chain to support possible replication of the search process and to allow for future research/researchers to build on the systematic review and update it at a later stage if desired. Only one of the systematic reviews appears to make some effort to deal with publication bias in some way by including evaluations conducted in a hand-full of languages.

Implementing these refinements, in conjunction with greater use of realist evaluation methods in underlying evaluations, could contribute towards the quality and explanatory power of evidence reviewed. More detailed theories of change could be developed and tested regarding how business development interventions impact on specific results. This would support greater explanation and understanding of causal dynamics behind the achievement or not of specific results, and greater utilisation of meta-reviews to inform strategy, policy and programme design.

The overall finding from the systematic reviews is that some positive impacts on various levels of outcomes were found, but that it is difficult to generalise these findings for a number of reasons. Three key weaknesses have been identified in the evaluations included in the systematic reviews:

1. The issue of how long outcomes can be expected after an intervention:
This is crucial issue as there are known to be numerous lags related to how businesses grow and develop over time and these need to be better understood. A key challenge for impact evaluation is both figuring out when to measure outcomes (Woolcock, 2009) as well as how often to measure outcomes. This was noted by Zandniapour et al. (2004) in their systematic review where the following was noted and the fact that this issue is repeatedly raised in the most recent systematic reviews (Grimm & Paffhausen, 2014; Mckenzie & Woodruff, 2012) suggest that there is a need for a systematic framework for assessing the impacts of business development interventions.

Time is an important correlate of impact. Depending on the type of intervention, different impacts manifest themselves at different points in time (e.g., trade projects need time for impacts to manifest themselves). A repeated theme in the studies was the need for more time to determine impacts on jobs, cost–benefit ratios and sustainability. One evaluation concluded that for these types of impacts, a time span of 5–10 years is needed (Study of Colombia and El Salvador Competitiveness Promotion Initiatives) (Zandniapour, 2014, p. 35).

Given the large interest of many governments in employment creation, studies that look only within a year or so of treatment ending may miss effects that take some time to be realised—or conversely, we may find that effects which look very promising in the short-term dissipate over time.

2. The absence of cost-effectiveness information.

Existing evidence seems to suffer from a method bias. Randomized Control Trials (RCTs) are applied particularly to small programmes, very poor areas and very specific target groups. This limits the generalisability of their findings. Quasi-experimental approaches have a wider scope of application than RCTs, but still the range of application is limited by the necessity to create a control group. Therefore, evaluations covering policies for which it is difficult, if not impossible, to apply (quasi-)experimental approaches and thereby link cause and effect in a rigorous way, will have no chance of being included in a systematic review of the type used here. However, for evaluations to give policy guidance, it would be desirable to draw on lessons learnt about the entire range of job programme designs. This leads to a kind of dilemma as it is impossible to include all types of programmes and meet the quality standards of a systematic review at the same time. For now, the way forward seems to be not to neglect the findings of other types of evaluations—excluded from systematic reviews—while being frank about their methodological shortcomings. In the long run, however, research might come up with methodological quality standards that cover those types of interventions not open to rigorous impact measurement yet.

3. Comparison of the effectiveness of entrepreneurship training which focuses on changing practices versus changing personalities:

Mckenzie & Woodruff (2012) note that most business training courses focus on particular business practices; however, there is a school of thought that the attitudes and personalities that business owners bring to the business are equally, if not more important. There is also a range of different training courses studied by psychologists which focus more on the personality of being an entrepreneur than on the specific skills (Glaub & Frese, 2011). Glaub, Frese, Fischer and Hoppe (2012) find some evidence to support a positive impact of such training in Uganda. While several studies have incorporated some aspect of aspirations or entrepreneurial attitudes into their content, to date there is no research which tests the relative contribution of both types of training.

4. Lack of disaggregated findings at a detailed intervention sub-type level:

The systematic review results do not disaggregate their findings to a level of detail regarding intervention sub-types that would be needed to be useful to most policy-makers and decision-makers. This limitation could be related to the number of evaluations that are available for analysis for specific sub-interventions. Over time, this should improve and therefore it will be possible to conduct systematic reviews that include interventions that are similar in most important aspects of design.

Limitations

This research has limited itself to a review of four systematic reviews primarily due to time constraints. For this reason, I did not contact the authors of the four systematic reviews to assist in identifying further reviews. In theory, it should be possible for future researchers to replicate and update this research and to add additional relevant systematic reviews to update this analysis over time.

I have not focused the review on one type of population group (e.g., youth, women, rural, urban etc.) as many of the interventions focus on more than one population group. I will, however, analyse findings on what works, or does not, for different population groups where these are contained in the examined systematic reviews.

A general challenge facing all systematic and meta-reviews is publication bias, where it has been found that statistically significant research is more likely to be published. It is not clear what solutions exist to deal with this problem (if any) and what the implications are for how the results of systematic and meta-reviews should be reported.

While this meta-review did include a mixed-methods synthesis, this synthesis was limited to the four systematic reviews being examined. It was beyond the scope of this meta-review to identify and include other literature related to broader effectiveness questions (e.g., how did the interventions work, under what circumstances, in what contexts etc.) although if time and resources were not constraints on this meta-review, such a broader scope and literature inclusion would have added considerable value to the findings of this meta-review.

The four systematic reviews with one exception, did not evaluate theories of change (presumably because these were not included in the evaluations under review) and this limited their analysis of the reasons why (significant) impacts may not have been identified in some cases. In other words, either implementation failures or design weaknesses common reasons why interventions may not have achieved impacts; however, it is difficult to identify these without a theory of change.

Concluding Observations

This meta-review has identified a range of challenges which exist regarding the review of evidence which has been conducted in the four systematic reviews of business development interventions under investigation. These include challenges of attribution as well as generalisability of findings.

A number of possible areas of improvement have been identified regarding the way in which systematic reviews are conducted. These include clearer purposes regarding influencing decision making, improvements in the ways that results of systematic reviews are reported (e.g., to focus findings on outcomes), and better documentation of search strategies and search logs so as to support the potential updating of systematic reviews with the latest research at some future stage.

In addition, improvements are needed in the way impact evaluations of business development interventions are conducted. Impact evaluations of business development interventions should ideally include

- panel data on outcomes from beneficiaries for the short, medium and long term to allow for distinctions to be made between short-, medium- and long-term changes in outcomes and an enhanced understanding of the processes of change which explain these changes over time.

- cost-effectiveness information to address the current serious gap from the perspective of decision-maker needs. It is important that intervention implementers provide the necessary numbers and researchers should go beyond the estimate of simple impacts, which is not really helpful for those who have to allocate resources across different interventions.

Three main areas warranting further research can be identified as follows:

1. There is a need for clear reporting requirements for systematic reviews of development interventions regarding how research methods and findings are reported. As Thomson et al. (2011, p. 207) note, ‘consistency in nomenclature would facilitate searching for and identification of this type of evidence. More generally, consistency in how methods and findings are reported would help to provide clarity for readers, as they would have a better sense of what an overview is, and is not, when using the results for decision making.’
2. Methodological quality standards are needed for a broader range of intervention types such as policies that are not amenable to experimental or quasi-experimental approaches. As Grimm & Paffhausen (2014) note, ‘Existing evidence seems to suffer from a method bias. RCTs are applied particularly to small programs, very poor areas and very specific target groups. This limits the generalizability of their findings. Quasi-experimental approaches have a wider scope of application than RCTs, but still the range of application is limited by the necessity to create a control group. Therefore, evaluations covering policies for which it is difficult, if not impossible, to apply (quasi-)experimental approaches and thereby link cause and effect in a rigorous way, will have no chance of being included in a systematic review of the type used here. However, for evaluations to give policy guidance, it would be desirable to draw on lessons learnt about the entire range of job program designs.’
3. There is a need for clearer frameworks for business development interventions, including typologies of interventions as well as definitions for key outcomes that can support stronger comparisons across impact evaluations. In addition, there is a need for a clear framework for reporting on job creation/destruction impacts. Despite job creation lying at the heart of many entrepreneurship and business development interventions, there does not seem to be any general framework or guidelines for monitoring job creation from government interventions. Such a framework would provide guidance on monitoring different types and quality of jobs at different time periods for different types of interventions.

Notes

1. http://betterevaluation.org/plan/synthesize_value/synthesize_across_evaluations
2. Atlas.ti stands for ‘Archiv fuer Technik, Lebenswelt und Alltagssprache’. Translated: archive for ‘technology, the life world and everyday language’. The extension ‘ti’ (pronounced TEE EYE) stands for text interpretation.

References

- Astbury, B., & Leeuw, F. L. (2010). Unpacking black boxes: mechanisms and theory building in evaluation. *American Journal of Evaluation*, 31(3), 363–381.
- Attanasio, O., Augsburg, B., de Haas, R., Fitzsimons, E., & Harmgart, H. (2011). *Group lending or individual lending: Evidence from a randomized field experiment in Mongolia*. London: European Bank for Reconstruction and Development.
- Augsburg, B., de Haas, R., Harmgart, H., & Meghir, C. (2012). *Microfinance at the margin: Experimental evidence*

- from *Bosnia and Herzegovina*. London: European Bank for Reconstruction and Development.
- Banerjee, A., Duflo, E., Glennerster, R., & Kinnan, C. (2014). *The miracle of microfinance? Evidence from randomized evaluation*. Cambridge: Massachusetts Institute of Technology.
- Bloom, N., Eifert, B., Mahajan, A., McKenzie, D., & Roberts, J. (2011). *Does management matter? Evidence from India*. Working Paper 16658. Cambridge, Massachusetts: National Bureau of Economic Research.
- Caird, J., Sutcliffe, K., Kwan, I., Dickson, K., & Thomas, J. (2015). Mediating policy-relevant evidence at speed: Are systematic reviews of systematic reviews a useful approach? *Evidence and Policy*, 11(1), 81–97.
- Cho, Y., & Honorati, M. (April 2013). *Entrepreneurship programs in developing countries*. Working Paper 6402. Washington, DC: World Bank Policy Research.
- Cho, Y., & Maddalena, H. (2013). *Entrepreneurship programs in developing countries a meta regression analysis*. Working Paper 6402. Washington, D.C: World Bank Policy Research.
- Crépon, B., Devoto, F., Duflo, E., & Parienté, W. (2014). *Estimating the impact of microcredit on those who take it up: Evidence from a randomized experiment in Morocco* (No. w20144). National Bureau of Economic Research.
- De Mel, S., McKenzie, D., & Woodruff, C. (2012). One-time transfers of cash or capital have long-lasting effects on microenterprises in Sri Lanka. *Science*, 335(6071), 962–966.
- Field, R., Pande, R., & Papp, J. (2010). Does microfinance repayment flexibility affect entrepreneurial behavior and loan default? *Microfinance Gateway*, 6.
- Glaub, M., & Frese, M. (2011). A critical review of the effects of entrepreneurship training in developing countries. *Enterprise Development and Microfinance*, 22(4), 335–353.
- Glaub, M., Frese, M., Fischer, S., & Hoppe, M. (2012). *A psychological personal Initiative training enhances business success of African business owners*. Mimeo, National University of Singapore Business School.
- Gough, D., & Elbourne, D. (2002). Systematic research synthesis to inform policy, practice and democratic debate. *Social Policy and Society*, 1(3), 225–236.
- Gough, D., Oliver, S., & Thomas, J. (2012a). *Introduction to systematic reviews*. London: SAGE Publications.
- Gough, D., Thomas, J., & Oliver, S. (2012b). Clarifying differences between review designs and methods. *Systematic Reviews*, 1(28), 1–9.
- Gough, D., Oliver, S., & Thomas, J. (2013). *Learning from research: Systematic reviews for informing policy decisions: A quick guide*. Working paper, The Alliance for Useful Evidence. London.
- Greenhalgh, T. (1997). How to read a paper: Papers that summarize other papers (systematic review and meta analyses). *British Medical Journal*, 315, 672–675.
- Grimm, M., & Paffhausen, A. L. (2014). Interventions for employment creation in micro, small and medium sized enterprises in low and middle income countries—a systematic review.
- International Labour Organization (2010). *How to build an enabling environment for youth entrepreneurship and sustainable enterprise*. Geneva.
- Karlan, D., & Zinman, J. (2010). Expanding credit access: Using randomized supply decisions to estimate the impacts. *Review of Financial Studies*, 23(1), 433–464.
- . (2011). Microcredit in theory and practice: Using randomized credit scoring for impact evaluation. *Science*, 332(6035), 1278–1284.
- Lavis, J. N. (2009). How can we support the use of systematic reviews in policymaking?. *PLoS Med*, 6(11).
- Mckenzie, D., & Woodruff, C. (2012). *What are we learning from business training and entrepreneurship evaluations around the developing world?* Working Paper 6202. Washington, D.C.: World Bank, Policy Research.
- OECD (2002). *Glossary of key terms in evaluation and results based management*. Evaluation and Aid Effectiveness, DAC Working Party on Aid Evaluation.
- Pawson, R., and Tilley, N. (1997). *Realist Evaluation*. London: SAGE Publications.
- Pawson, R. (2004). *Realist synthesis: An introduction*. Leeds: University of Leeds.
- . (2006). *Evidence-based policy: A realist perspective*. London: SAGE Publications.
- . (2013). *The science of evaluation: A realist manifesto*. London: SAGE Publications.
- Pawson, R., Greenhalgh, T., Harvey, G., & Walshe, K. (2005). Realist review: A new method of systematic review designed for complex policy interventions. *Journal of Health Services Research and Policy*, 10(1), 21–34.

- Pawson, R., & Tilley, N. (1997). *Realist evaluation*. London: SAGE Publications.
- Petticrew, M., & Roberts, H. (2006). *Systematic reviews in the social sciences: A practical guide*. Oxford: Blackwell Publishing.
- Rogers, P. J. (2007). Theory-based evaluation: Reflections ten years on. In S. Mathison (Ed.), *New directions for program evaluation* (pp. 63–67). San Francisco, CA: Jossey-Bass.
- Thomson, D., Russell, K., Becker, L., Klassen, T., & Hartling, L. (2011). The evolution of a new publication type: Steps and challenges of producing overviews of reviews. *Research Synthesis Methods, 1*(3–4), 198–211.
- Waddington, H., White, H., Snilstveit, B., Hombrados, J. G., Vojtkova, M., Davies, P., ... & Valentine, J. C. (2012). How to do a good systematic review of effects in international development: A tool kit. *Journal of development effectiveness, 4*(3), 359–387.
- Weiss, C. H. (1997). *Theory-based evaluation: Past, present, and future. New directions for evaluation*. San Francisco, CA: Jossey-Bass.
- Wildschut, L. (2014). *Theory-based evaluation, logic modelling and the experience of South African non-governmental organisations* (Doctoral dissertation). South Africa: University of Stellenbosch.
- Woolcock, M. (2009). Toward a plurality of methods in project evaluation: A contextualised approach to understanding impact trajectories and efficacy. *Journal of Development Effectiveness, 1*(1), 1–14.
- World Bank. (2013). *Framing the global landscape of entrepreneurship education and training programs*. Washington, D.C.
- Zandniapour, L., Sebstad, J., & Snodgrass, D. (2004). *Review of evaluations of selected enterprise development projects. Microenterprise report, 3*, Washington D.C: USAID.