

Vocational and business training to increase women's participation in higher skilled occupations in low- and middle-income countries: protocol for a systematic review

Marjorie Chinen,¹ Thomas de Hoop,¹ María Balarin,² Lorena Alcázar²

Submitted to the Coordinating Group of:

<input type="checkbox"/>	Crime and Justice
<input type="checkbox"/>	Education
<input type="checkbox"/>	Disability
<input checked="" type="checkbox"/>	International Development
<input type="checkbox"/>	Nutrition
<input type="checkbox"/>	Social Welfare
<input type="checkbox"/>	Other:

Plans to co-register:

<input checked="" type="checkbox"/>	No		
<input type="checkbox"/>	Yes	<input type="checkbox"/> Cochrane	<input type="checkbox"/> Other
<input type="checkbox"/>	Maybe		

Date Submitted: 31 March 2016

Date Revision Submitted: 28 July 2016

Approval Date:

Publication Date: 16 December 2016

¹ American Institutes for Research (AIR)

² Group for the Analysis of Development (GRADE)

BACKGROUND

The Problem

Although women's employment possibilities have improved with the rise of globalization, women in low- and middle-income countries tend to be overrepresented in informal labour markets, work in precarious conditions, receive lower salaries than men, and have few opportunities for learning and advancement (Borges Månsson & Färnsveden, 2012). Duflo (2012) reports that "women are less likely to work, they earn less than men for similar work, and are more likely to be in poverty even when they work" (p. 1052). Women often perform jobs that have low skill requirements and frequently work in occupations that are highly feminized, tend to be less socially valued, and pay lower wages (Aedo & Walker, 2012; Altman, 2006; International Labour Organization [ILO], 2015; ILO, 2016).

A recent ILO report has documented the limited opportunities for women in the labour market (ILO, 2016). The report shows that women face higher unemployment and underemployment than men, are more often employed in the informal labour market and in family enterprises, and are overrepresented in lower skill sectors. For example, a greater proportion of women are employed in the services sector (61.5 per cent versus 42.6 per cent of men), where women are particularly overrepresented in feminized positions such as "clerical, services, and sales" and "elementary" occupations (ILO, 2016). Moreover, although men and women face equal rates of wage and salaried employment (around 52 per cent), men are more likely to own their own business than women (3.7 per cent of men versus 1.4 per cent of women). Meanwhile, and while data on informal sector employment is scant, women are 'believed to constitute most of the informal workforce in the developing world' (UNGEI 2012). Work in the informal sector is characterised by low pay and low productivity (ILO, 2016). Women working in informal employment do not gain access to social protection, such as pensions, and this may contribute to the fact that 71.8 per cent of all employed women do not have any type of maternity protection (ILO, 2016).

A range of factors contribute to the high proportion of unemployed and underemployed women. These factors include cultural norms regarding the place of women in employment, the role of women in domestic and care work, and the lack of adequate job market opportunities. Women all over the world spend a disproportionate amount of time doing domestic and care work. This time commitment is even higher in low- and middle-income countries, where the division of domestic labor often follows traditional patterns and women assume most, if not all family responsibilities (ILO, 2009). Women may also have a preference for jobs that are compatible with their domestic responsibilities, such as part-time and flexible jobs, both of which are scarce in low- and middle-income countries (ILO, 2009). However, even when part-time and flexible job opportunities exist, allowing women to combine work and family responsibilities, there is evidence that these jobs do not constitute 'a path to decent work' (ILO, 2005). This lack of opportunities contributes to the choice of women to remain self-employed in small-scale enterprises or in domestic and care

work, as opposed to pursuing professional careers in higher skilled occupations (ILO, 2009; ILO, 2012).

Evidence also suggests that lack of pre-service and in-service training opportunities leads many women to seek employment in low-skilled areas and in highly feminized occupations (ILO, 2016). The opportunities for women in other occupations are limited by the domestic obligations and limited skill sets of many women in developing countries (Katz, 2008; Mansson & Farnsveden, 2012). As a result, many young women seek employment in the informal sector, leaving them only with low-skilled employment characterized by minimum income potential, long working hours, and unequal power relationships, which may lead to exploitation. Limited data are available on informal labour markets in low- and middle-income countries. Nonetheless, evidence suggests that women constitute the large majority of the informal workforce in developing countries (UNGEI 2012).

The Intervention

National governments and development agencies have created a range of vocational and business training programs that aim to increase the participation of women in higher skilled and more secure, formal occupations. Such programs typically focus on improving the skills of women, foster entrepreneurship to expand employment, increase earning opportunities; and ultimately, reducing poverty (Blattman & Ralston, 2015; McKenzie & Woodruff, 2012), as discussed in more detail below.

Vocational training programs often target low-income, unemployed or underemployed individuals who have already left the formal schooling system, but they can also be part of the formal education system. These programs typically include preparing participants for jobs that are related to a specific occupation or trade through the acquisition of knowledge, practical cognitive and non-cognitive skills, and changing attitudes. Trainings include courses in administrative occupations (e.g., marketing, secretarial work, sales), manual occupations (e.g., electrician, cooking assistant), or fairly skilled occupations (e.g., account assistant, information technology (IT) specialist).

Many vocational training programs also combine the development of specific occupational skills with strategies to facilitate access to job opportunities, for instance, through internships, on-the-job training or by actively connecting participants with potential employers. For example, a representative program named “Jovenes en Accion” from Colombia provided three months of in-classroom and three months of on-the-job training to young people between the ages of 18 and 25 in the two lowest socioeconomic strata of the population (Attanasio, Kugler, & Meghir, 2008). Similarly, the Peruvian youth labour training program named “PROJoven” offered poor youths three months of training and a three-month internship with a local firm for beneficiaries who successfully completed their coursework (Ñopo, Robles, & Saavedra, 2007).

Additionally, some vocational training programs include the development of life skills among program participants. Program content ranges from specific job related life skills, such as preparing a *Curriculum Vitae* and interpersonal relationships to broader skills, such as those related to reproductive health and household economics. For example, Gap Inc's Personal Advancement and Career Enhancement program focuses on improving non-cognitive skills through learning modules that emphasize communication, problem solving and decision making, and time and stress management, among other skills (Gap Inc, 2015).

Business training programs provide active support to the poor to try to improve the performance of small and medium enterprises around the world (McKenzie & Woodruff, 2013). These programs attempt to offer skills in the form of business management training and assistance to small and medium enterprises. Business training may be offered by governments, microfinance organizations and non-governmental organizations. The length of these programs range from as short as two days to several months, and they are commonly offered to groups, although some programs also provide additional one-on-one follow-up training.

The majority of these trainings focus on general business skills that could be applied to most businesses (e.g., keeping business records and encouraging small business owners to separate household and business finances), while fewer focus on technical knowledge or sector-specific content or on trying to change entrepreneurial attitudes or aspirations. Several programs focus specifically on increasing women's understanding of the value chain for their product, sources of raw material and access to markets. For example, a PRASAC program in Cambodia helped weavers improve their businesses by identifying and building relations with suppliers, wholesalers, retailers, and designers (GTZ, 2003).

Finally, some business training programs are offered with additional incentives. These incentives could include asset or in-kind transfers, such as livestock or inventory, capital as a small cash loan or grant, or saving accounts. For example, a program implemented by BRAC Bangladesh named the Targeting the Ultra Poor (TUP) program offered a variety of business activities that ranged from livestock rearing to small retail operations to very poor and rural woman in selected rural communities. The program was combined with complementary and intensive asset-specific training and regular follow-up visits by asset specialists (Bandeira, Burgess, Gulesci, Rasul, & Sulaiman, 2013).

Both vocational and business training programs typically target low-income, unemployed or underemployed individuals, but there are also a range of programs that focus explicitly on building the skills of women. For instance the Jordan New Opportunities for Women program focuses on providing opportunities to young, female college graduates in Jordan to find work (Groh, Krishnan, McKenzie, & Vishwanath, 2012), while the Women for Women International Social Protection and Cash Transfer program in Nigeria offers biweekly meetings with business, vocational and life-skills training, including lessons in health awareness, decision

making, negotiation, and civic participation (Mcilvaine & Oser, 2014). Both programs focus on women exclusively.

Vocational and business training programs that target women in particular may also include additional components to address barriers to women's participation in either the training itself or broader barriers to their participation in the labour market. Such components may include the provision of basic stipends to women to participate in the training and additional support for women with children (such as child care facilities or stipends). They may also include negotiating with women's families for permission to participate in the training (Ñopo, Robles, & Saavedra, 2007; Revenga, Riboud and Tan, 1994; Aedo and Nuñez 2004).

How the Intervention Might Work

Vocational or business training programs aim to improve outcomes such as women's labour market outcomes, income and empowerment by developing women's knowledge, skills and opportunities, and by doing so increasing women's participation in higher-skilled occupations (Revenga & Shetty, 2012). We developed a theory of change for both vocational training and business training programs, outlining how these programs might work in improving outcomes. They map out the causal chain between interventions, outputs, intermediate outcomes, and final outcomes, as well as the assumptions underlying the theory of change (White, 2009). Using such program theories help provide a framework for the analysis in the review and determine relevant outcomes for the systematic review. We outline the theories of change for each program area below.

Theory of Change for Vocational Training Programs Targeting Women

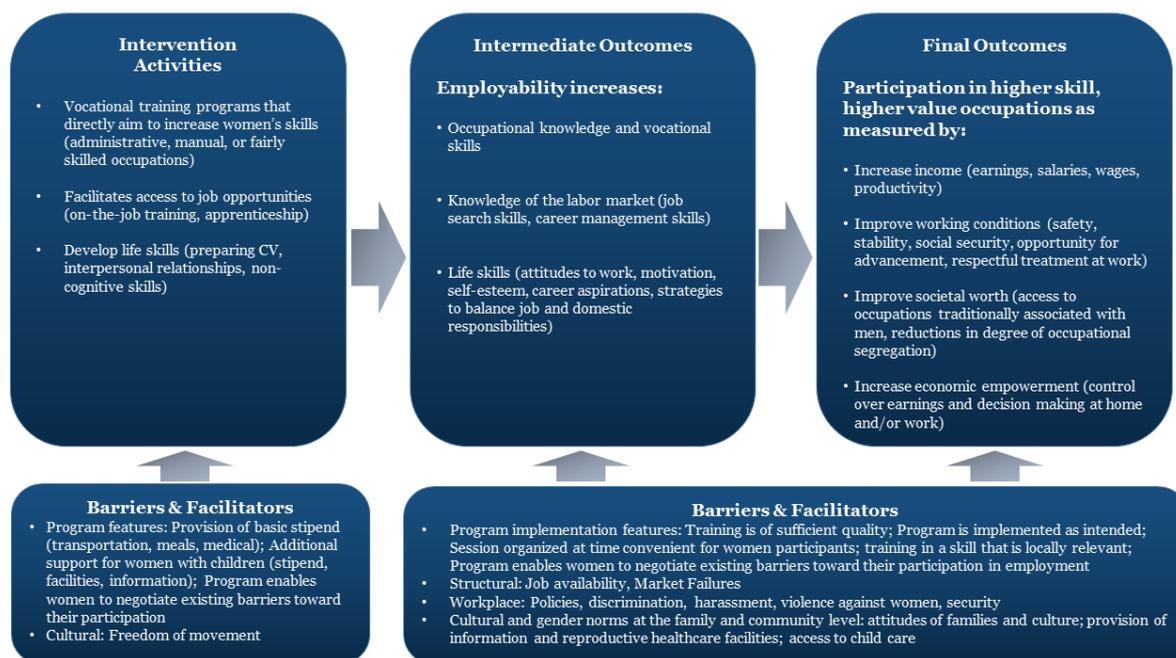
Figure 1 provides a theory of change for how vocational training programs may improve women's access to higher skilled, higher valued occupations and improve final outcomes such as income and empowerment. First, vocational programs can provide training that directly aims to increase women's skills (e.g., marketing, sales, account assistant, IT specialist), facilitate access to job opportunities (e.g., by on-the-job training or apprenticeship), or equip women with "life skills" to improve aspects of life such as interpersonal relationships and reproductive health.³ Second, vocational training can result in improvements in women's employability by increasing women's occupational knowledge and vocational skills, women's knowledge of the labour market (e.g., job search skills, career management skills), and women's life skills such as attitudes toward work, motivation, self-esteem, career aspirations, and strategies to balance job and domestic responsibilities.

In turn, these improvements in women's skills and employability may result in increases in women's participation in higher skilled, higher valued occupations. Access to higher skilled, higher valued occupations can in turn increase women's income and productivity; improve

³ We will not include studies that exclusively focus on life skills training to improve reproductive health. Every study that we include needs to focus on either vocational or business training.

women’s working conditions (such as safety, stability, social security) and opportunities for advancement; and reduce the degree of occupational segregation by improving access to occupations traditionally dominated by men.⁴ Finally, the theory of change indicates that access to higher skilled, higher valued occupations may lead to an improvement in women’s economic empowerment, such as women’s control over earnings and decision making at home and at work.

Figure 1. Theory of Change for Vocational Programs Targeting Women



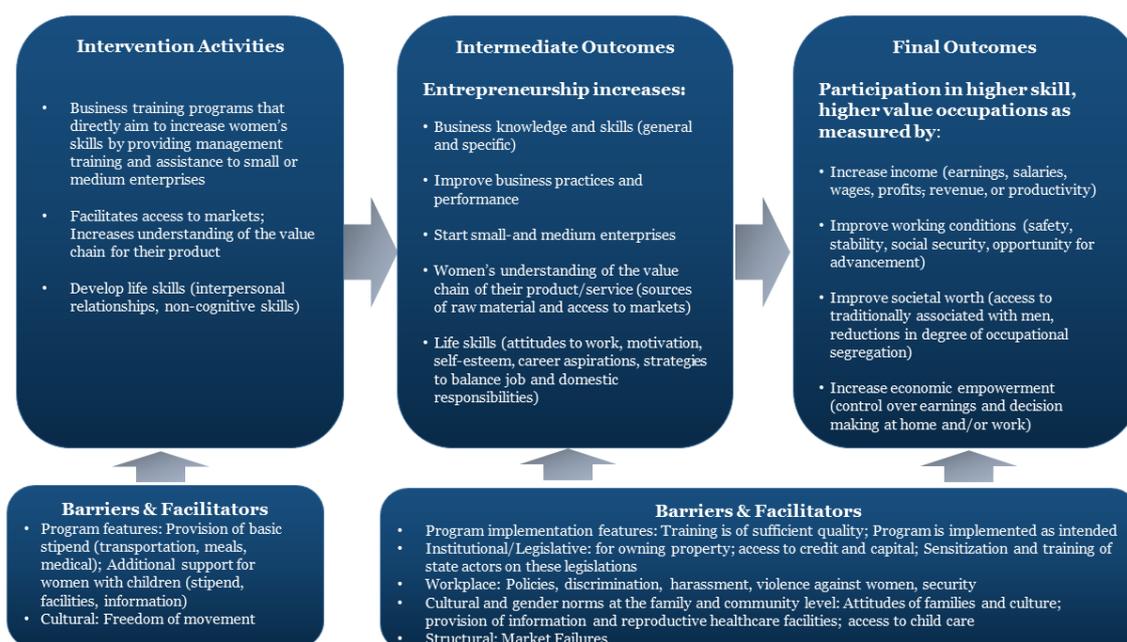
Theory of Change for Business Training Programs Targeting Women

Figure 2 provides a theory of change for how business programs can contribute to increasing women’s participation in higher skilled, higher valued occupations, and by doing so improve final outcomes such as income and empowerment. Business training programs typically provide training to equip women with business and life skills; provide assistance to small or medium enterprises; provide capital in the form of assets, cash, loans, or saving accounts and facilitate access to markets. Providing these inputs can result in improvements in a range of intermediate outcomes associated with entrepreneurship, such as business knowledge and skills (generic business skills and specific skills, such as knowledge about value chains), and their “life skills” such as attitudes toward work, motivation, self-esteem, career aspirations, and strategies to balance job and domestic responsibilities. These changes may in turn increase the likelihood that women start their own business, and/ or improve their business performance through the application of better business practices. Further,

⁴ In the case of the impact evaluation of the Peruvian youth training program named “PROJoven,” females increased their participation in the occupational groups of sales personnel, restaurant and food service workers, and domestic workers, which were traditionally filled by males.

these improvements can result in increases in income (as measured by earnings, salaries, wages, profits, revenue, or productivity), improvements in working conditions (stability, social security, opportunity for advancement) and reduce the degree of occupational segregation by improving access to occupations traditionally dominated by men. Finally, changes in intermediate and final outcomes could lead to an improvement in women’s economic empowerment, such as women’s control over earnings and decision making at home and at work.

Figure 2. Theory of Change for Business Programs Targeting Women



Barriers and facilitators: The validity of the theory of change relies on a number of assumptions. If these assumptions do not hold the causal pathways in the theories of change may break down. One key assumption is that women are able to participate in training programs. But several barriers may limit women’s participation in vocational or business training. Women require freedom of movement to participate in either vocational or business trainings, however restrictions on women’s movement outside the house or on their participation in trainings by other household members may prevent this. Similarly, women may spend too much time on domestic chores or care work to allow them to participate in trainings, particularly if there is no access to child care or additional stipends for child care.

Other critical assumptions are related to the quality of the training, legislative conditions, and access to credit. If the training is of insufficient quality it may not develop women’s knowledge, skills or entrepreneurship. Effective training requires high-quality trainers and relevant curricula with tailored content to teach women vocational or business skills. Moreover, legislative conditions may need to be in place to facilitate gains in women’s participation in higher-skilled occupations. For example, women often require access to

formal credit to benefit from business training because without access to credit, it may be challenging to set up a small business.⁵ Even if women are able to participate in training and develop their skills, women with children may require access to affordable child care facilities for the training to translate into participation in higher-skilled occupations. Finally, policies in the workplace, including those relating to discrimination, harassment and job security may also influence the extent to which programs succeed in changing final outcomes.

The barriers and facilitators are linked to various key stakeholders and different types of interventions may be effective depending on the barrier(s) that are most relevant in a specific context. For example, if discrimination in the workplace is a major barrier, vocational training may not be effective if it does not address women's ability to negotiate discrimination against women in the workplace.

A variety of market failures could be linked to the funding of vocational and business training programs. These market failures are associated with the question "if individuals can achieve high returns to participation in vocational and business training, why do these individuals not participate in these programs instead of waiting for the government to provide them?". Market failures that prevent participation in the absence of government-funded training could include missing credit markets, information failures, and insurance market failures (Almeida et al., 2012). Missing credit markets could lead to an underinvestment of trainees in business training because of a lack of access to credit. Information failures could lead to a lack of knowledge about the availability of vocational or business training, which could in turn result in a lower participation in these training. Insurance market failures may lead to underinvestment in training when the returns to vocational and business training are uncertain (Almeida et al., 2012).

Why It Is Important to Do the Review

Relevance for policy and practice

The effectiveness of these types of programs for women may be of interest to a range of decision makers. Perhaps most importantly, governments in developing countries make significant investments in vocational and business training programs. For example, many Ministries of Labour in Latin-America, such as Peru, Argentina, and Colombia, have invested in the ProJoven program (Aedo and Nuñez, 2004; Ñopo, Robles, & Saavedra, 2007; Ministerio de Trabajo, 2016).

Moreover, the findings of this review will also be relevant for developing policies and programs to achieve two of the sustainable development goals (SDGs) in particular. SDG 5 calls for gender equality and empowering all women and girls (SDG 5), while SDG 8 calls for

⁵ At the same time, however, changes in legislation may not be sufficient if bank staff remains unconvinced about the ability of women to repay loans, even in the presence of legislation that protects women's rights.

sustained, inclusive and sustainable economic growth, full and productive employment and decent work for all (Sustainable Development Knowledge platform, n.d.).

Several development organizations highlight the importance of stimulating women to participate in higher-skilled occupations. For example, the Economic Commission for Latin America and the Caribbean, the Food and Agriculture Organization, UN Women, UNDP, and the ILO highlight the importance of investing in development programs to improve employment access and quality for women in Latin America and the Caribbean (ECLAC, FAO, UN Women, UNDP, & ILO 2013). In addition, the United Nations Conference on Trade and Development (2016) argues that better education and on-the-job training are required to let trade create opportunities for women's empowerment and well-being. These initiatives demonstrate the importance policy makers attach to increasing the employment of women in higher skilled, higher valued occupations. To achieve this goal, L&MIC governments and development agencies have created a wide range of programs, such as vocational and business training, which aim to improve the skills of women (Blattman & Ralston, 2015; McKenzie & Woodruff, 2012).

Finally, conversations with key stakeholders from UN Women, the Canadian Department of Foreign Affairs, Trade and Development, and the Inter-American Development Bank suggest a demand for evidence on what works to improve women's participation in higher-skilled occupations. This systematic review aims to respond to this demand. A systematic review will also be important for policymakers and practitioners who would like to learn about ways to improve the design of vocational and business training programs.

Review of existing literature

A preliminary mapping of evidence identified over 20 experimental and quasi-experimental studies of the effectiveness of vocational and business training in stimulating women's employment in higher skilled jobs (Chinen et al., 2015), as well as a number of studies of the structural barriers associated with cultural gender norms and the different positions of men and women in the labour market in low- and middle-income countries.

Several authors have reviewed this, and related literature. Tripney et al. (2013) provide a systematic review of evidence on the effects of technical and vocational education training programs on employment outcomes, concluding that such programs improve outcomes overall. However, the review did not differentiate the effects by gender, nor include any qualitative studies. Moreover, Chinen et al. (2015) found an increase in the number of randomized controlled trials since the completion of the review of Tripney et al. (2013).

More recent non-systematic literature reviews have been critical of the effectiveness of vocational and business training programs in improving outcomes, raising concerns about the size of the effect of such programs (Blattman & Ralston, 2015; McKenzie & Woodruff, 2012). Blattman & Ralston (2015) suggest that skills training only has limited effects on poverty or stability. McKenzie & Woodruff (2012) find that few studies with an emphasis on

business training find evidence for positive effects on profits or sales one year after the start of the program. Other systematic and non-systematic reviews and meta-analyses that review related literature include reviews that focus on the effects of business support services on job creation, labour productivity and the ability of firms to invest (Piza et al., 2016), the impact of programs targeted at micro-entrepreneurs on job creation (Grimm & Paffhausen, 2014), the effects of entrepreneurship development interventions on women entrepreneurs (Patel, 2014), and the impact of active labor market program evaluations (Card, Kluve, & Weber, 2015).

We did not identify any systematic review of this literature focusing on the effects of vocational and business training on outcomes for women. Therefore this study will be the first to assess the effect of vocational and business training on women's socioeconomic outcomes. For this purpose, we will synthesise evidence from quantitative studies of vocational and business training programs. We will also review a broader range of literature to assess the barriers and facilitators of the effectiveness of such programs. This analysis will be of particular relevance for policymakers and practitioners who wish to improve the design of vocational and business training programs.

OBJECTIVES

The primary objective of this systematic review is to synthesize the evidence on the effects of vocational and business training programs on women's participation in higher skilled, higher valued occupations and other-labour market outcomes, including employment, income, working conditions, societal worth, and economic empowerment. The secondary objective is to improve our understanding of the barriers to, and facilitators of, the effectiveness of vocational and business training for women and how they operate.⁶ To achieve these goals will address the following research questions:

Primary Research Question

1. What are the effects of vocational and business training programs on women's participation in higher skilled, higher valued occupations, women's income, working conditions, societal worth, and economic empowerment?

Secondary Research Question

2. What are the barriers to, and facilitators of, the effectiveness of vocational and business training programs?
3. What are the program design and implementation characteristics associated with effective vocational and business training programs for women?

⁶ The influence of cultural and gender norms will vary across different regions. Such regional variations will be taken into consideration in the quantitative and the qualitative analyses by relying on random effect meta-analyses and a narrative review that takes contextual characteristics into consideration.

METHODOLOGY

Criteria for Including and Excluding Studies

Types of Study Designs

The primary research question on the effectiveness of interventions (question 1) will be addressed using quantitative experimental or quasi-experimental studies. We will include both experimental studies that use random assignment to the intervention and quasi-experimental designs with non-random assignment. To be included, quasi-experimental studies need to use either known allocation rules (such as assignment on the basis of a threshold on a continuous variable or geographic variation in the assignment of the program) or include pre-and post-test measures of the outcome variable of interest. Knowledge about allocation rules may enable the use of regression discontinuity designs or natural experiments to determine the impact of the program, while the inclusion of pre-test and post-test measures of outcome variables will enable researchers to use methods that control for selection-bias such as interrupted time series models, difference-in-difference regression analysis, statistical matching (for example propensity score matching or covariate matching), instrumental variables, or Heckman selection models.

We will only include only quasi-experimental studies that use methods that can credibly address selection-bias. Cook et al. (2008) and Shadish (2011) demonstrate that quasi-experimental studies can address concerns with selection-bias, but only under certain conditions. Controlling for selection-bias from covariates and including pre-test measures of outcome variables are particularly effective in reducing selection bias (Steiner et al., 2010; Shadish, 2011). We will therefore only include quasi-experimental studies that either include a baseline measurement of the outcome of interest or other relevant confounding factors or where allocation rules enable the use of regression discontinuity designs or analyses on the basis of natural experiments. We will exclude quasi-experimental studies that do not include a baseline measure of the outcome of interest and do not enable the use of either regression discontinuity designs or analyses on the basis of natural experiments.

To address the secondary research questions (questions 2 and 3), we will include a broader range of evidence with relevant details on intervention design, implementation and context. To include these studies and documents we will use a similar approach as Snilstveit et al. (2015) in their systematic review regarding the effects of education programs. In line with that review, included studies and documents will focus on one of the interventions that was evaluated in the primary research question, and have to meet at least one of the following three criteria:

1. A qualitative study collecting primary data using qualitative methods of data collection and analysis, and reporting some information on all of the following: 1) the research question, 2) procedures for collecting data and 3) sampling and recruitment.
2. A process evaluation assessing whether the intervention is being implemented as intended and how it is working. Process evaluations may include the collection of qualitative and quantitative data from different stakeholders to cover subjective issues, such as perceptions of intervention success or more objective issues, such as how an intervention was operationalized.
3. A project document describing the background, design and implementation features of an intervention. This document does not typically include much primary evidence, but it should provide factual information about interventions to be considered. The purpose of including project documents in our review is to ensure we have sufficient information about the context and interventions in included studies.

Types of Participants and Settings

We will include studies that focus on interventions that include women who are 18 years or older in low- and middle-income countries, as defined by the World Bank.⁷ In cases where the intervention participants are not exclusively women, studies will be eligible only if effects on women are assessed separately from those on men. We will include studies about the effects of vocational and business training regardless of the employment status or skill level of women at the time of the intervention.

Types of Interventions

The interventions included in this review will be vocational and business training programs that aim to increase the level of skills of the disadvantaged, unemployed, or underemployed; foster entrepreneurship to expand employment; and increase income prospects of women.

To identify interventions that focus on high skill occupations for women we take Acemoglu & Aedo's (2011) classification of occupations according to skill requirements as a starting point. They define three broad occupational categories:

1. Low Skills occupations, that include: routine manual occupations (that require working at a pace determined by speed of equipment; spending time making repetitive motions); as well as routine cognitive occupations (that require repeating the same tasks; being exact or accurate).
2. Medium Skills occupations that include; non-routine manual physical occupations (such as operating vehicles, mechanized devices, or equipment); as well as non-

⁷ We are using the 2014 LMICs definition, which includes Argentina, Hungary, Seychelles, and Venezuela. These countries, however, were categorised as high-income countries in the July 2015 update.

routine cognitive interpersonal occupations (which require establishing and maintaining personal relationships; guiding, directing, and motivating subordinates; coaching and developing others.)

3. High Skills occupations, which the authors describe as those occupations requiring the performance of non-routine cognitive analytical tasks (such as analysing data or information; thinking creatively, interpreting information for others).

We slightly adjust this definition of high skill occupations for the purpose of this review. While the definition of Acemoglu & Aedo (2011) is useful, we can expect a degree of variation as to what constitutes a high, medium or low-skill occupation in the context of low and middle-income countries. For example, a bank teller would constitute a low-value, low-skill, routine cognitive job in a high-income country, but in a low- or middle-income country, a job as a bank teller would qualify as a high-skill occupation. For this reason, we will include studies that focus on programs which provide women with either high skills or medium skills as defined by Acemoglu & Aedo's framework.

Vocational training programs could also include one or more of the following components (in addition to vocational training):

- Strategies to facilitate access to job opportunities (through internships, on-the-job training or by actively connecting participants with potential employers).
- Development of life skills among program participants (specific job related ones such as preparing a *Curriculum Vitae* or broader skills such as those related to reproductive health and household economics).

We will also include studies that focus on business training programs providing management training and assistance to small and medium enterprises. In addition to business training, these programs can include one or more of the following components:

- Strategies to facilitate access to markets and increase understanding of the value chain for the enterprise product.
- Development of life skills among program participants (specific job related ones such as preparing a *Curriculum Vitae* or broader skills such as those related to reproductive health and household economics).

We will only include evaluations of vocational or business training programs that include at least one of the components discussed above.

Excluded interventions

Several other criteria are required for interventions to be included in the review. We will exclude vocational and business training programs targeted exclusively at men. Interventions that train women to work in low-skill occupations fall outside the scope of this review. Thus, we will exclude studies that focus on interventions that aim to increase the labour market participation of women in low-skill occupations. We used a scoping review to

further understand how low-skill occupations are defined within low- and middle-income countries. On the basis of the scoping review, we decided to exclude studies of programs that increase women's access to the domestic service or the agricultural labour market because our scoping review indicated that these sectors almost exclusively provide employment in routine manual jobs. We will also exclude those studies in this systematic review.

Some vocational training programs are offered within high schools or tertiary institutions. We will not include studies that focus on the latter types of vocational training programs because these programs are likely to influence outcomes through different mechanisms and will thus require a different theory of change.

Types of Comparison Conditions

Eligible comparison conditions will include no intervention, pipeline, or "business as usual." In those studies that we include to address the secondary research questions, a comparison condition will not be necessary for the document to be included.

Types of Outcome Measures

To address research question 1 we will include studies that focus on a range of intermediate and/or final outcomes. Studies need to assess at least one of the outcomes outlined below.

Intermediate outcomes:

Knowledge and Skills: Training programs aim to improve participants' knowledge and skills. We will include any studies that measure women's business knowledge and skills; women's understanding of the value chain of their product/service; women's occupational knowledge and vocational skills and women's life skills. Studies that measure these outcomes through either tests, administrative or self-reported data will all be included.

Business Practices: We will include any measure of the adoption of business practices that is associated with good business outcomes. These practices may include regular bookkeeping, keeping clear records, good planning, and other practices that are taught during business trainings. We will include studies that measure the adoption of these and other good business practices (as reported by the authors of the study) through surveys.

Employment (status, occupation): We include all studies that focus on a range of measures of women's employment status and type of occupation, including whether women are employed, the status of their employment (full-time/part-time, permanent/temporary, formal/ informal) and the type of occupation they are employed in.

Final outcomes:

Income (Earnings, Wages, or Salaries): We include measures of income that women earn through paid work, defined as earnings, wages, or salaries. We will include studies that measure earnings or wages through administrative data as well as self-reported data in

which women are asked about their earnings, wages, or salaries. We will include only studies that measure these incomes at the individual level to ensure that we can distinguish between income of men and women. We will not include studies that proxy for income by relying on expenditure or other consumption data unless those data are able to distinguish between the expenditure levels of men and women.

Revenue or Profits: We define revenue or profits as the total amount of income generated by the sale of goods or services from a business and the total amount of income generated by the sale of goods or services from a business minus the costs, respectively. We will include studies that measure revenue or profits through administrative and self-reported survey data. We will include only studies that report the revenues or profits of women or women-owned businesses specifically.

Working Conditions: Working conditions are comprised of “a broad range of topics and issues, from working time... to remuneration, as well as the physical conditions and mental demands that exist in the workplace” (ILO, n.d.). More specifically, working conditions may include factors such as safety, stability of employment, social security, opportunities for advancement; job permanence; career and employment security; health and well-being at the workplace; and work–life balance; as well as respectful treatment at work. We will include all studies that measure these factors through administrative or self-reported data.

Societal Worth: We define improved societal worth as women’s access to occupations traditionally associated with men, which in turn might lead to a reduction in the degree of occupational segregation. Occupational segregation and the feminization of certain occupations may contribute to gender inequalities in earnings and in the societal value attributed to women’s work (Borges, Mansson, & Farnsveden, 2012). Thus, improved societal worth may improve the contribution of women on the labour market. We will include studies that measure societal worth or occupational segregation through survey or administrative data.

Economic Empowerment: In line with Brody et al. (2015), we define economic empowerment as “the ability of women to access, own, and control resources.” Women’s empowerment can be measured through survey data, for example, by asking women about their ownership of assets and land, division of domestic labour across men and women, and control over financial decision making (Brody et al., 2015). In contrast to Brody et al. (2015), we, however, do not consider degree of women’s participation in paid employment as an economic empowerment indicator because we emphasize the importance of high-skilled labour.

Studies addressing questions 2 and 3: Outcome measures will not be considered to filter qualitative studies, which will serve to address the secondary research questions. From these documents we will include descriptive information about programme design and implementation, context and resources, as well as any findings addressing questions 2 and 3

on barriers and facilitators of intervention success or failure.

Other criteria for inclusion/ exclusion

We will consider studies published in English or Spanish after 1990. We will not exclude studies based on publication status.

Search Strategy

We developed a comprehensive search strategy in consultation with information specialists. Our search strategy will enable us to identify published and unpublished literature by focusing on relevant academic and institutional databases, citation tracking, and snowballing of references. For searches of electronic databases we will use a detailed search string, whereas searches of grey literature and institutional websites will typically rely on more simple, tailored searches.

The search strings are designed to return studies that include at least one keyword in the following four themes:

1. Participants: woman, female, wife, mother, gender, occupational stratification, empowerment businesswoman, woman industry.
2. Employment: employment, employability, employer, employee, work, workforce, job, vocation, career, occupation, livelihood, workplace, part-time, casual, informal, wages, labour market, school-to-work.
3. Intervention: skill, upskill, life skills, training, retraining, mentor, apprentice, internship, prestige, expertise, professional, qualification, education, coaching, leadership, entrepreneur, human capital, business management, business support, capacity building.
4. Setting: transitional, low-income, middle-income, third world, developing, underdeveloped, underserved nations/countries/populations/economies; LMIC, LAMI, low GDP, low GNP.⁸

To capture both quantitative and qualitative literature relevant to the research questions, we will not include search strings for study design, comparison condition, and outcome measures. Using these criteria in the search strategy would exclude relevant qualitative studies in addition to quantitative and mixed-methods studies that omit this information from the title and abstract.

Appendix 1 provides an example of a search string. Each search string will be adapted to fit the syntax of the database host in order to utilize Boolean operators (AND/OR), wildcards, truncation, and other database search features. Below we list the sources covered by our search.

⁸ Low- and middle-income countries are identified using the World Bank 2014 definition.

Electronic Sources

Comprehensive database searches will include the following paid-access and free-access electronic databases:

Paid-Access Databases

1. Web of Science
2. Scopus
3. ASSIA
4. IBSS
5. PAIS
6. ERIC
7. Econlit
8. Academic Search Complete
9. Business Source Corporate
10. Gender Studies Database

Open-Access Databases

1. Ideas-REPEC
2. Campbell Library
3. 3ie Impact Evaluation Repository
4. Labourdoc

Hand search: Selected Journals

We will hand search the most recent issues of the journals listed below for any papers not yet indexed in electronic databases.

1. American Economic Review
2. Quarterly Journal of Economics
3. American Economic Journal: Applied Economics
4. Journal of Labour Economics
5. Journal of Development Economics
6. Journal of Development Effectiveness
7. Economic Development & Cultural Change
8. International Labour Review
9. World Bank Economic review
10. World Development
11. Journal of Development Studies
12. Journal of Human Resources
13. Journal of Education and Work
14. Journal of International Women's Studies
15. Gender & Development
16. Feminist Economics
17. Feminist Review

Grey Literature

Grey literature searches will include a review of institutional websites and research funders. In addition to the following institutional websites, key papers will be examined in search of other relevant institutional sources.

Institutional Websites and Research Funders

18. International Labor Organization (<http://www.ilo.org>)
19. World Bank (<http://www.worldbank.org>)
20. ELDIS (<http://www.eldis.org>)
21. DfID R4D (<http://r4d.dfid.gov.uk>)
22. ASEAN (<http://www.asean.org>)
23. IADB (<http://www.iadb.org>)
24. UNGEI (<http://www.ungei.org>)
25. UN-Women (<http://www.unwomen.org>)
26. J-PAL (<http://www.povertyactionlab.org>)
27. IPA (<http://www.poverty-action.org>)
28. British Library for Development Studies (<http://blds.ids.ac.uk/>)
29. World Bank DIME (<http://www.worldbank.org/en/research/dime>)
30. World Bank enGENDER Impact
(<http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTGENDER/0..contentMDK:23457844~pagePK:210058~piPK:210062~theSitePK:336868,00.html>)
31. World Bank Policy Research Working Paper Series
<http://documents.worldbank.org/curated/en/docsearch/document-type/620265>
32. NBER (National Bureau of Economic Research): <http://www.nber.org/>
33. Social Science Research Network (SSRN):
<http://papers.ssrn.com/sol3/DisplayAbstractSearch.cfm>
34. IZA Discussion Papers
<http://www.iza.org/en/webcontent/publications/papers>
35. Centre for Economic Policy Research Discussion Papers
<http://cepr.org/content/discussion-papers>
36. BREAD working papers
<http://ibread.org/bread/papers>

Citation tracking

Forward and backward snowballing of the references of key papers will provide additional studies for review that may not be found in database searches. Citation searches will be conducted in Google Scholar, Scopus, and Web of Science. This set of key papers will include: anchor papers identified by the authors (listed below), key papers identified by an advisory group of reviewers of the scoping study we conducted by Chinen et al. (2015), key papers identified by Advisory Group, key papers identified by external funder (3ie), studies that pass the inclusion criteria and additional key papers identified from institutional websites.

Table 1 presents the list of key papers identified by authors and reviewers to be used in citation searches.

Table 1. Key Papers

Study
Attanasio, O., Kugler, A., & Meghir, C. (2008). <i>Training disadvantaged youth in Latin America: Evidence from a randomized trial.</i>
Bandiera, O., Burgess, R., Gulesci, S., Rasul, I., & Sulaiman, M. (2013). <i>Capital, skills and the economic lives of the poor: Recent evidence from field experiments.</i>
Blattman, C., & Ralston, L. (2015). <i>Generating employment in poor and fragile states: Evidence for labour market and entrepreneurship programs.</i>
Blattman, C., Fiala, N., & Martinez, S. (2013). <i>Generating skilled self-employment in developing countries: Experimental evidence from Uganda.</i>
Borges Månsson, A., & Färnsveden, U. (2012). <i>Gender and skills development: A review.</i>
Bravo, D., Sanhueza, C., & Urzúa, S. (2008). <i>An experimental study of labour market discrimination: Gender, social class and neighborhood in Chile.</i>
Buvinic, M., & Furst Nichols, R. (2014). <i>Promoting women's economic empowerment: What works?</i> World Bank.
Chigateri, S., Jugran, T., Nanda, R., & Nandi, R. (2015). "Learning from STEP Evaluations" in <i>Engendering meta-evaluations: Towards women's empowerment.</i>
Cho, Y., & Honorati, M. (2013). <i>Entrepreneurship programs in developing countries: A meta regression analysis.</i>
Erulkar, A., Bruce, J., Dondo, A., Sebstad, J., & Matheka, J. (2006). <i>Tap and Reposition Youth (TRY): Providing social support savings and microcredit opportunities for young women in areas with high HIV prevalence.</i>
Ibarrarán, P., & David Rosas, S. (2008). <i>Evaluating the impact of job training programs in Latin America: Evidence from IDB-funded operations.</i>
International Centre for Research on Women (2003). <i>Equal opportunity and youth employment: A background report for the Equal Opportunity Working Group of the Youth Employment Network.</i>
Kabeer, N. (2005). <i>Is microfinance a "magic bullet" for women's empowerment? Analysis of findings from South Asia.</i>
Kabeer, N. (2012). <i>Women's economic empowerment and inclusive growth: Labour markets and enterprise development.</i>
Katz, E. (2008). <i>Programs promoting young women's employment: What works?</i> The Adolescent Girls Initiative.
McKenzie, D., & Woodruff, C. (2012). <i>What are we learning from business training and entrepreneurship evaluations around the developing world?</i>
International Center for Research on Women. (2012). <i>Catalyzing Growth in the Women-Run Small and Medium Enterprises Sector (SMEs).</i>
Nanda, P., Mishra, A., Walia, S., Sharma, S., Weiss, E., & Abrahamson, J. (2013). <i>Advancing women, changing lives: An evaluation of Gap Inc.'s PACE program.</i>
Ñopo, H., Robles, M., & Saavedra, J. (2008). <i>Occupational training to reduce gender segregation: The impacts of PROJoven.</i>
Sudarshan, R. M., & Bhattacharya, S. (2009). <i>Through the magnifying glass: Women's work and labour force participation in urban Delhi.</i>

Study

Tripney, J., Hombrados, J., Newman, M., Hovish, K., Brown, C., Steinka-Fry, K., et al. (2013). *Technical and vocational education and training (TVET) interventions to improve employability and employment of young people in low- and middle-income countries: A systematic review.*

Targeted Search for Addressing Questions 2 and 3

After we have identified studies for inclusion to answer review question 1, we will undertake targeted searching for qualitative studies, as well as process evaluations, implementation documents and other relevant documents for those interventions evaluated in the included studies. We will conduct citation tracking of included studies to identify any relevant papers and conduct internet and database searches using the names of programs from included studies. We will also search databases of project documents and websites of implementing agencies to identify project documents. The targeted search will involve the following steps:

1. **Advisory group:** contact advisory group members from UN Women, the Canadian Department of Foreign Affairs, Trade and Development, the Inter-American Development Bank, 3ie, the World Bank, BRAC and other key stakeholders to request project documentation about vocational and business training programs that were evaluated under the first research question.
2. **Citation tracking:** conduct forward-and backward citation tracking of included studies to identify relevant papers and other documents.
3. **Search by program name:** implement searches for the program name on the internet and in specific databases. We will also conduct searches using Boolean logic through combining “program name” and country/funder/implementer. We will screen the first 50 hits on google as in Snilstveit et al. (2015).
4. **Targeted searches of funder and implementer websites:** conduct searches on the websites of implementers and funding agencies. We will again conduct searches using Boolean logic through combining “program name” and country/funder/implementer as well.

Screening

Screening will take place in two phases: first on the basis of titles and abstracts and then on the basis of full texts.

Screening Phase 1

Titles and abstracts of the search results of database, grey-literature, and citation searches will be uploaded to the EPPI-Reviewer 4 online system (<http://eppi.ioe.ac.uk/eppireviewer4>). Duplicate studies will be identified and removed automatically using the duplicates feature built into EPPI-Reviewer 4. The inclusion criteria will be piloted by two reviewers; a sample of the studies will be independently assessed for

relevance by each reviewer (see Appendix 2). Screening results will be compared, and the inclusion criteria will be adapted as needed. All titles and abstracts will then be screened against the inclusion criteria outlined above, retaining potentially relevant studies for screening at full text (phase 2).

Screening Phase 2

In the second phase, we will review the full text of all studies that pass Phase 1 screening. Multiple reviewers will independently assess studies against the inclusion criteria and studies meeting all criteria will be retained for inclusion in the review.

Studies that pass Phase 2, will be screened for the secondary research question based on the exclusion/inclusion criteria outlined above. Therefore, studies may pertain to only research question 1, or to both research questions 1 and 2. For example, a study that only uses quantitative methods may pertain only to research question 1. While a study that uses mixed methods –with a focus on measuring the impact of vocational training on women’s labour market outcomes and on the barriers or facilitators that played a role- may pertain to both research questions 1 and 2.

Critical appraisal

Quantitative studies addressing question 1: Risk of Bias Assessment

We will determine the rigor of the quantitative studies using an adaptation of a set of criteria, to assess risk of bias in experimental and quasi-experimental studies (Hombrados & Waddington, 2012). Two researchers with expertise in experimental and quasi-experimental methods will assess the risk of the following biases:

1. Selection bias and confounding, based on quality of identification strategy to determine causal effects and assessment of equivalence across the beneficiaries and non-beneficiaries.
2. Performance bias, based on the extent of spill-overs to women in comparison groups and contamination of the control or comparison group.
3. Outcome and analysis reporting biases, including:
 - a. The use of potentially endogenous control variables
 - b. Failure to report non-significant results
 - c. Other unusual methods of analysis
4. Other biases, including:
 - a. Motivation and courtesy biases (Hawthorn effect and John Henry effect)
 - b. Coherence of results
 - c. Retrospective baseline data collection
 - d. Differential attrition bias and overall level of attrition
 - e. Other biases, such as strong researcher involvement in the implementation of the intervention.

The risk of bias assessment will not be used to exclude studies. Instead, we will rely on the risk of bias assessment to determine whether studies can be considered low, medium, or high risk of bias in the four risk of bias domains (selection bias and confounding, performance bias, outcome and analysis reporting bias, and other biases). Appendix 5 presents the risk of bias assessment tool.

Studies addressing questions 2 and 3: Critical Appraisal

To address questions 2 and 3, we will include different types of documents as outlined in the inclusion criteria. Critical appraisal of qualitative studies, process evaluations and project documents (including a broader range of evidence) can be complicated particularly as there is a lack of existing tools and criteria for quality (Noyes et al., 2011). Following Snilstveit et al. 2015, we adopt different approaches to appraise these types of studies/ documents, as outlined below.

We will assess the quality of the first group, included qualitative studies, using and adaptation of the nine-item Qualitative Research Checklist developed by the Critical Appraisal Skills Programme (CASP), presented in Table 2 below, making judgments on the adequacy of stated aims, the data collection methods, the analysis, the ethical considerations, and the conclusions drawn (Critical Appraisal Skills Programme, 2013). For each item, two researchers will determine whether the study had adequately met the item and gave “yes,” “no,” or “can’t tell” responses. If researchers disagree, they will discuss the item to reach consensus. We will rate studies that score 0–2 “no” or “can’t tell” responses as with minor limitations, studies that score 3–5 “no” or “can’t tell” responses as with medium limitations, and studies that score 6–9 “no” or “can’t tell” responses as major limitations.

In the case of process evaluations, although there are no commonly used critical appraisal tools, assessment of methods of data collection should be considered. To achieve this goal, we propose to use the adapted version of the CASP checklist presented in Table 2.

Finally, project documents are typically descriptive and do not include much analysis of primary evidence. Therefore we will not formally appraise the quality of such documents. Rather we will assess the relevance of the documents for understanding program design and implementation and assess if they provide factual information about interventions. If several such sources are available we will extract data from all sources for purposes of triangulation.

Table 2. Quality Appraisal Criteria

Criteria	Coding
<i>Screening Question:</i> Is there a clear statement of study aims of the research?	Yes / Can't tell / No
<i>Screening Question:</i> Is a qualitative methodology appropriate?	Yes / Can't tell / No
Was the research design appropriate to address the aims of the research?	Yes / Can't tell / No
Was the recruitment strategy appropriate to address the aims of the research?	Yes / Can't tell / No
Were the data collected in a way that addressed the research question?	Yes / Can't tell / No
Has the relationship between researcher and participants been adequately considered?	Yes / Can't tell / No
Have ethical issues been taken into consideration?	Yes / Can't tell / No
Was the data analysis sufficiently rigorous?	Yes / Can't tell / No
Is there a clear statement of findings?	Yes / Can't tell / No

Data Extraction

Two team members with expertise in quantitative research will work independently to extract information from each quantitative study included in the review. The data will include details of study design, intervention, comparison group, outcome, and data for effect-size calculation (more details below). Similarly, two team members with expertise in qualitative research will extract data from studies included to address questions 2 and 3. Data to be extracted will include descriptive details about intervention design, and more analytical findings on process, implementation and context. The draft data extraction forms are provided in Appendix 3 and 4. Any disagreements will be resolved through discussion.

Measures of Treatment Effects

We will extract information from each quantitative study to allow for the estimation of standardized effect sizes, and associated standard errors and 95 per cent confidence intervals. Where possible, we will use natural units, such as percentage changes in income and percentage point increases in employment and synthesize studies using those effect sizes. However, this approach will only be feasible when studies use the same outcome measures, which is not likely to be the case for all studies.

In all other cases, we will calculate two types of effect sizes. We will calculate the Hedges' *g* sample-size-corrected standardized mean differences (SMDs) for continuous outcome variables, which measure the effect size in units of standard deviation of the outcome variable. We will calculate odds ratios (ORs) for binary outcome variables.

We will calculate standardized mean differences (Cohen's d) by dividing the mean difference with the pooled standard deviation by applying the formula in Equation 1:

$$SMD = \frac{Y_t - Y_c}{S_p} \quad (1)$$

SMD refers to the standardized mean differences, Y_t refers to the outcome for the treatment group, Y_c refers to the outcome for the comparison group, and S_p refers to the pooled standard deviation.

The pooled standard deviation S_p can be calculated by relying on the formulas in Equations 2 and 3:

$$S_p = \frac{\sqrt{((SD_y)^2 * (n_t + n_c - 2)) - \frac{\beta^2 * (n_t * n_c)}{n_t + n_c}}}{n_t + n_c} \quad (2)$$

$$S_p = \frac{\sqrt{(n_t - 1) * s_t^2 + (n_c - 1) * s_c^2}}{n_t + n_c - 2} \quad (3)$$

We will use Equation 2 for regression studies with a continuous dependent variable. In this equation, SD_y refers to the standard deviation for the point estimate from the regression, n_t refers to the sample size for the treatment group, n_c refers to the sample size for the control group, and β refers to the point estimate. We will use Equation 3 when there is information about the standard deviation for the treatment group and the control group.

We will correct the standardized mean difference for small sample size bias by relying on formula 4, which transforms Cohen's d to Hedges' g.

$$SMD_{corrected} = SMD_{uncorrected} * \left(1 - \frac{3}{4 * (n_t + n_c - 2) - 1}\right) \quad (4)$$

We will rely on Equation 5 to estimate the standard error of the standardized mean difference:

$$SE = \sqrt{\frac{n_t + n_c}{n_c * n_t} + \frac{SMD^2}{2 * (n_c + n_t)}} \quad (5)$$

Where possible, we will calculate odd ratios by relying on 2 by 2 contingency tables (Lipsey & Wilson, 2001), as described in Figure 3.

Figure 3. Estimation of Odds-Ratios

	Frequencies	
	Success	Failure
Beneficiaries	A	B
Comparison Group	C	D

We will calculate the odds-ratio using Equation 6, where \overline{ES} refers to the effect size:

$$\overline{ES} = \frac{ad}{bc} \quad (6)$$

Where studies use linear probability models, we will assume linearity in the estimation of standardized effect sizes as in Brody et al. (2015). This assumption will allow us to use the same 2 by 2 contingency table on the basis of the baseline means and impact estimates from the linear probability model. For example, if we observe a mean baseline value for the comparison group of 0.097 and an effect size of 5.1 percentage points, then we will assume that the follow-up value for the treatment group would be $0.097+0.051=0.148$, and we will assume that the follow-up value for the comparison group will be 0.097.

We will use adjusted standard errors for those studies that use outcome variables that are clustered at a higher level of aggregation than the individual or household level but do not take this into consideration in the estimation of the standard errors and confidence intervals. For these studies, we will apply corrections to the standard errors and confidence intervals using the variance inflation factor (Higgins & Green, 2011):

$$SE_{corrected} = SE_{uncorrected} \times \sqrt{(1 + (m - 1) \times ICC)} \quad (7)$$

Here, m is the number of observations per cluster, and ICC is the intraclass correlation coefficient. We will estimate an intraclass correlation for each of the outcome measures from data of a study that can be considered representative for our sample of included quantitative studies and for which we are able to obtain data on the outcome measures. We will use a range of plausible ICCs to examine the effects of these ranges on the outcome of our synthesis.

Missing Data

In cases where it is not feasible to estimate the effect size because of missing information, we will contact the authors of the primary studies to request the missing information that is required to calculate the effect sizes. In those cases in which we are not able to retrieve missing data, we will extract or impute effect sizes and associated standard errors based on commonly reported statistics such as the t or F statistic or exact p or z-values using David Wilson's practical meta-analysis effect-size calculator. Where studies do not report sample sizes for the treatment and the control or comparison group, we will assume equal sample sizes across the groups.

Methods for Handling Dependent Effect Sizes

We will include only one effect size per study in any single meta-analysis.⁹ Where studies report more than one effect size on the basis of different statistical methods, we will select the effect size with the lowest risk of bias. We will use the effect size from the most recent publication when several publications reporting on the same study report different effect sizes. We will check working paper versions for data on additional outcomes. Where studies present several impact estimates for different outcomes variables that measure the same

⁹ We will use robust variance estimation (Hedges, Tipton, & Johnson, 2010) where that is feasible, but we do not expect to include a sufficient number of studies.

construct, we will use a sample-size weighted average to create a “synthetic effect size.” In cases where more than one study use the same data set to measure an outcome variable, we will extract the effect size from the study with the lowest risk of bias. In cases where one study measures the same outcome variable at different periods in time, we will extract data on all time periods, but include the outcome measure that is closest to the time period of the measurement in other studies included in the same meta-analysis. Estimates from other time points will be discussed narratively. Finally, if studies include more than one treatment arm, we will include the effect size from the treatment arm that is most similar to the other programs that are included in the meta-analysis as recommended in Borenstein et al. (2009).

Methods of synthesis

Review question 1: Statistical analysis

If meta-analysis is feasible, we will address question 1 by synthesising studies using an inverse-variance, random effects model due to the anticipated heterogeneity in the included studies. We will only combine estimates from studies assessing the effect of comparable programs on similar outcome constructs (Wilson et al., 2011). Thus we will conduct separate meta-analyses for vocational and business training programs, running different analyses for different outcome constructs.

We will examine the heterogeneity of the effect sizes visually and by estimating the I-squared and Q as well as tau-squared (Borenstein et al., 2009). If we have sufficient studies we will investigate factors explaining heterogeneity by using inverse-variance weighted meta-regressions and stratified meta-analysis according to the following contextual and methodological moderator variables: (1) type of intervention component, (2) intervention components, (3) geography, (4) age of beneficiaries, (5) type of treatment effect estimate (eg: intention to treat, treatment effect on the treated etc.) (6) time to follow-up.

We will also examine the relative effectiveness of studies that evaluate different types of interventions through stratified meta-analysis. For this purpose we will use a similar approach as Fu et al. (2011), and Brody et al. (2015) who compare point estimates and then test differences in effect sizes between programs with different intervention constructs formally if the overlap in effect-sizes is minor or non-existent.

We will perform sensitivity analysis for two methodological effect size moderators:

- Risk of bias status for each risk of bias category.
- Study design (randomized controlled trials versus quasi-experimental studies).

We will start our analysis with separate meta-analyses of randomized controlled trials and quasi-experimental evaluations for determining the effects of vocational and business training. Then we will use an iterative approach to determine the potential bias from pooling

randomized controlled trials and quasi-experimental evaluations and studies with low, medium, and high risk of bias for each of the types of bias we assessed in our risk of bias assessment. We will present the results of each meta-analysis, highlighting any differences in results between studies with different levels of risk of bias. This approach is in line with the approach used in a recent systematic review on the effects of women's self-help groups on women's empowerment (Brody et al., 2015).

Publication Bias

We will use various methods to determine the potential for publication bias. First, we will assess the potential for publication bias using funnel plots. Second, we will conduct Egger's test to determine the potential for publication bias in studies that focus on vocational or business training.

Qualitative synthesis

Techniques for integrating, comparing and synthesising qualitative literature have been variously described as "meta-ethnography" (Harden, 2010), "narrative synthesis" (Snilstveit, Oliver, & Vojtkova, 2012), and, more recently, as "meta-synthesis" (Sandelowski, Docherty, & Emden, 1997; Walsh and Downe 2005; Zimmer 2006).

In spite of the different terms, authors agree that unlike what happens in quantitative meta-analysis, where the aim is to aggregate the findings from multiple studies, in the case of qualitative synthesis, the main aim is not "aggregation" but rather "interpretation." Rather than pooling findings from studies, key concepts are translated within and across studies, and the synthesis product is a new interpretation (Harden, 2010).

A specific challenge faced in this process is the highly contextual nature of qualitative findings, which means that findings or interpretations from a study cannot be merely extracted and compared but need to be considered closely to the context in which they were generated. Equally challenging is the fact that qualitative interpretations are even more mediated than quantitative findings by the epistemological and theoretical perspectives used by the researchers. All of this means that qualitative synthesis processes need to endeavour strongly to the characterization of included studies before moving on to develop middle-range and generalizing theories about their findings (Zimmer, 2006).

Thus the synthesis of included qualitative studies will proceed along the following stages:

1. **Characterization:** During this stage, we will code and describe included studies along various items, including methodology (types of methods used, sample characteristics), epistemological perspective, theoretical framing, context, and thematic focus. This phase is important in order to appropriately locate studies within their empirical, theoretical, and epistemological context.

2. Thematic coding: During this stage, we will code findings from included studies for themes and concepts.
3. Summarizing and comparison: Summaries of the research findings will be elaborated and compared in order to identify similarities and differences. This comparison will be carried out taking into account the initial characterization of the studies, so that themes and concepts are appropriately read taking into account contextual nuances and the theoretical perspective used in the different studies.
4. Qualitative meta-synthesis: Based on the characterization, coding and summarizing stages in the included study findings will be synthesized into broader explanatory theories. These theories will serve as a basis to examine the influence of structural barriers and facilitators and gender norms on the effectiveness of vocational and business training. Furthermore, we will examine the importance of contextual characteristics and how these are linked to the structural barriers and facilitators and gender norms.

Integrated synthesis (review questions 1, 2 and 3)

Once we have completed the syntheses above we will provide an integrated synthesis of the findings. We will use the program theories provided above to present the findings from the different syntheses with the aim of providing an integrated narrative synthesis addressing the objectives of the review. We will use a summary of findings tables following the GRADE (Schünemann et al., 2011) and CerQual (Lewin et al., 2015) approaches to facilitate transparent and systematic presentation of our findings.

We will develop a series of matrices to identify the features of interventions and contexts that appear to act as barriers and facilitators of the interventions' effectiveness. We will initially conduct the cross-case analysis by intervention type, but we will also attempt to identify any overarching themes across intervention types. Finally, we will provide an integrated synthesis of the findings addressing all review questions.

REFERENCES

Key Papers for Citation Searches

- Attanasio, O., Kugler, A., & Meghir, C. (2008). *Training disadvantaged youth in Latin America: Evidence from a randomized trial*. London, England: EDePo Centre for the Evaluation of Development Policies. Retrieved from <http://www.nber.org/papers/w13931>
- Bandiera, O., Burgess, R., Gulesci, S., Rasul, I., & Sulaiman, M. (2013). *Capital skills and the economic lives of the poor: Recent evidence from field experiments*. Retrieved from http://www.womeneconroadmap.org/sites/default/files/Bandiera%20et%20al_Capital%20Skills%20Economic%20Lives%20of%20Poor.pdf
- Blattman, C., Fiala, N., & Martinez, S. (2013). Generating skilled self-employment in developing countries: Experimental evidence from Uganda. *Quarterly Journal of Economics*. Retrieved from <http://dx.doi.org/10.2139/ssrn.2268552>

- Blattman, C., & Ralston, L. (2015). *Generating employment in poor and fragile states: Evidence for labour market and entrepreneurship programs*. Retrieved from <http://dx.doi.org/10.2139/ssrn.2622220>
- Borges Månsson, A., & Färnsveden, U. (2012). *Gender and skills development: A review* (Background Paper for the EFA Global Monitoring Report 2012). New York, NY: United Nations Girls' Education Initiative. Retrieved from http://www.ungei.org/files/UNGEI_Gender_and_Skills_Print_v1_3.pdf
- Bravo, D., Sanhueza, C., & Urzúa, S. (2008). *An experimental study of labour market discrimination: Gender, social class and neighborhood in Chile*. Washington, DC: Inter-American Development Bank. Retrieved from <http://dx.doi.org/10.2139/ssrn.1815907>
- Buvinic, M., & Furst Nichols, R. (2014). *Promoting women's economic empowerment: What works?* Washington, DC: World Bank. Retrieved from <https://openknowledge.worldbank.org/bitstream/handle/10986/20629/WPS7087.pdf>
- Chigateri, S., Jugran, T., Nanda, R., & Nandi, R. (2015). "Learning from STEP Evaluations" in *Engendering meta-evaluations: Towards women's empowerment*. New Delhi, India: Institute of Social Studies Trust. Retrieved from http://www.academia.edu/20321542/Engendering_Meta-Evaluations_Towards_Womens_Empowerment
- Cho, Y., & Honorati, M. (2013). *Entrepreneurship programs in developing countries: A meta regression analysis*. Washington, DC: World Bank. Retrieved from <http://dx.doi.org/10.1596/1813-9450-6402>
- Critical Appraisal Skills Programme (CASP) (2013). *Collabouration for Qualitative Methodologies*. Available from <http://www.casp-uk.net/wp-content/uploads/2011/11/CASP-Qualitative-Research-Checklist-31.05.13.pdf>.
- Erulkar, A., Bruce, J., Dondo, A., Sebstad, J., & Matheka, J. (2006). *Tap and Reposition Youth (TRY): Providing social support savings and microcredit opportunities for young women in areas with high HIV prevalence*. New York, NY: Population Council. Retrieved from http://www.popcouncil.org/uploads/pdfs/TABriefs/PGY_Brief15_TRY.pdf
- Evans, K. (1995). *Barriers to participation of women in technological education and the role of distance education*. The Commonwealth of Learning. <http://hdl.handle.net/11599/465>
- Harden, A. (2010). "Mixed-methods systematic reviews: integrating quantitative and qualitative findings." The National Center for the Dissemination of Disability Research (NCDDR), Technical Brief, No. 25.
- Higgins, J. P. T. & Green, S., eds. (2011). *Cochrane Handbook for Systematic Reviews of Interventions*. Version 5.1.0 (updated March 2011). The Cochrane Collabouration Available at <http://www.cochrane-handbook.org/>.
- Ibarrarán, P., & David Rosas, S. (2008). *Evaluating the impact of job training programs in Latin America: Evidence from IDB-funded operations*. Washington, DC: Inter-American Development Bank. Retrieved from http://www.iza.org/conference_files/ELMPDC2009/ibarraran_p4263.pdf

- International Center for Research on Women (2012). *Catalyzing growth in the women-run small and medium enterprises sector (SMEs)*. Washington, DC: International Center for Research on Women. Retrieved from <http://www.icrw.org/sites/default/files/publications/Catalyzing-growth-in-the-women-run-small-and-medium-enterprises-sector.pdf>
- Kabeer, N. (2005). Is microfinance a “magic bullet” for women’s empowerment? Analysis of findings from South Asia. *Economic and Political Weekly*, 4709–4718. Retrieved from <http://www.jstor.org/stable/4417357>
- Kabeer, N. (2012). *Women’s economic empowerment and inclusive growth: Labour markets and enterprise development*. Ottawa, ON, Canada: International Development Research Centre. Retrieved from <https://www.soas.ac.uk/cdpr/publications/papers/file80432.pdf>
- Katz, E. (2008). *Programs promoting young women’s employment: What works?* (Launch of Adolescent Girls Initiative). Washington, DC: World Bank. Retrieved from <http://siteresources.worldbank.org/INTGENDER/Resources/GenderYouthEmploymentKatz.pdf>
- Mansson, A.B., & Farnsveden, U. (2012). Gender and skills development: A review (Background paper for the EFA Global Monitoring Report 2012). New York, NY: United Nations Girls’ Education Initiative. Retrieved from http://www.ungei.org/files/UNGEI_Gender_and_Skills_Print_v1_3.pdf
- Mcilvaine, K., & Oser, C. (2014). Confidence, capacity-building, and cash: How a holistic program for ultra-poor women in Rwanda and the Democratic Republic of Congo yielded sustained impact for graduates. Retrieved from <https://www.ids.ac.uk/files/dmfile/Graduationconferencepaper-McilvaineandOser.pdf>
- McKenzie, D., & Woodruff, C. (2012). *What are we learning from business training and entrepreneurship evaluations around the developing world?* Bonn, Germany: Institute for the Study of Labour. Retrieved from <http://ftp.iza.org/dp6895.pdf>
- Nanda, P., Mishra, A., Walia, S., Sharma, S., Weiss, E., & Abrahamson, J. (2013). *Advancing women, changing lives: An evaluation of Gap Inc.’s PACE program*. Washington, DC: International Center for Research on Women. Retrieved from <http://www.icrw.org/files/publications/Equal-Opportunity-and-Youth-Employment.pdf>
- Ñopo, H., Robles, M., & Saavedra, J. (2008). Occupational training to reduce gender segregation: The impacts of PROJoven. *Economía*, 31(62), 33–54. Retrieved from <http://revistas.pucp.edu.pe/index.php/economia/article/download/1177/1136>
- Riche, M. F. (2003). *Equal opportunity and youth employment: A background report for the Equal Opportunity Working Group of the Youth Employment Network* (Working Paper Series). Washington, DC: International Center for Research on Women. Retrieved from <http://www.icrw.org/files/publications/Equal-Opportunity-and-Youth-Employment.pdf>
- Sudarshan, R. M., & Bhattacharya, S. (2009). Through the magnifying glass: Women's work and labour force participation in urban Delhi. *Economic and Political Weekly*, 59–66. Retrieved from <http://www.jstor.org/stable/25663838>

Tripney, J., Hombrados, J., Newman, M., Hovish, K., Brown, C., Steinka-Fry, K., et al. (2013). *Technical and vocational education and training (TVET) interventions to improve employability and employment of young people in low- and middle-income countries: A systematic review*. Oslo, Norway: The Campbell Collaboration. Retrieved from <http://www.campbellcollaboration.org/lib/project/227/>

Additional References

- Acemoglu, D., & Autor, D. (2011). Skills, Tasks, and technologies: Implications for employment and earnings. *Handbook of Labour Economics* 4(B), 1043–171.
- Aedo, C. & S. Nuñez. (2004). *The impact of training policies in Latin America and the Caribbean: the case of Programa Joven*. Inter-American Development Bank. <http://www.iadb.org/res/publications/pubfiles/pubr-483.pdf>
- Aedo, C., & Walker, I. (2012). *Skills for the 21st Century in Latin America and the Caribbean*. Washington, DC: World Bank. DOI: 10.1596/978-0-8213-8971-3
- Almeida, R., Hirshleifer, S., McKenzie, D., Ridao-Cano, C., & Yener, A.L. (2012). The impact of vocational training for the unemployed in Turkey: Pre-analysis plan. Retrieved from https://www.povertyactionlab.org/sites/default/files/documents/ISKURIE_AnalysisPlan_v4.pdf
- Altman, M. (2006). Wage determination in South Africa: What do we know? *Transformation: Critical Perspectives on Southern Africa*, 60(1), 58–89.
- Borenstein, M., Hedges, L.V., Higgins, J.P.T., and Rothstein, H. (2009). *Introduction to Meta-Analysis*. Chichester: Wiley.
- Brody, C., de Hoop, T., Vojtkova, M., Warnock, R., Dunbar, M., Murthy, P., et al. (2015). *Economic self-help group programs for improving women's empowerment: A systematic review*. Oslo, Norway: The Campbell Collaboration.
- Card, D., Kluge, J., & Weber, A. (2015). *What works? A meta analysis of recent active labor market programs*. IZA Discussion Paper 9236. Retrieved from <http://ftp.iza.org/dp9236.pdf>
- Chinen, M., de Hoop, T., Balarin, M., Alcazar, L., Sennett, J., & Mezarina, J. (2015). *What works in increasing women's participation in higher skilled, higher valued occupations in low and middle income countries?* (A Scoping Review).
- Cook, T.D., Shadish, W.R., & Wong, V.C. (2008). The conditions under which experimental & observational studies produce comparable causal estimates: New findings from within-study comparisons. *Journal of Policy Analysis & Management*, 27(4): 724-750.
- Duflo, E. (2012). Women empowerment and economic development. *Journal of Economic Literature*, 50(4), 1051–1079. Retrieved from <http://dx.doi.org/10.1257/jel.50.4.1051>
- Duvendack, M., Palmer-Jones, R., & Vaessen, J. (2014). Meta-analysis of the impact of microcredit on women's control over household decisions: Methodological issues and substantive findings. *Journal of Development Effectiveness*, 6(2), 73–96. Retrieved from <http://www.tandfonline.com/doi/abs/10.1080/19439342.2014.903289>

- ECLAC, FAO, UN Women, UNDP, & ILO (2013). Decent work and gender equality: Policies to improve employment access and quality for women in Latin America and the Caribbean. Retrieved from http://www.oitcenterfor.org/sites/default/files/file_publicacion/wcms_229430_3.pdf
- Frantzi, K., Ananiadou, S., & Mima, H. (2000). Automatic recognition of multi-word terms. *International Journal of Digital Libraries* 3(2), 117–132. Retrieved from <http://link.springer.com/article/10.1007/s007999900023>
- Fu, R., Gartlehner, G., Grant, M., Shamliyan, T., Sedrakyan, A., Wilt, T.J., Griffith, L., Oremus, M., Raina, P., Ismaila, A., Santaguida, P., Lau, J., Trikalinos, T.A. (2011). Conducting qualitative synthesis when comparing medical interventions: AHRQ and the effective health care program. *Journal of Clinical Epidemiology*, 64: 1187-97.
- Gap Inc. (2015). *Advancing Women to Advance the World: Program Overview*. Retrieved from <https://www.bewhatspossible.com/DevAssets/Docs/PACEProgramOverview2015.pdf>
- Grimm, M., Paffhauzen, A.L. (2014). *Do interventions targeted at micro-entrepreneurs and small and medium-sized firms create jobs? A systematic review of the evidence for low and middle income countries*. IZA Discussion Paper, (8193). Retrieved from <http://ftp.iza.org/dp8193.pdf>
- Groh, M., Krishnan, N., McKenzie, D. J., & Vishwanath, T. (2012). *Soft skills or hard cash? The impact of training and wage subsidy programs on female youth employment in Jordan*. World Bank Policy Research Working Paper, (6141). <http://dx.doi.org/10.1596/1813-9450-6141>
- GTZ. (2003) “Guide to Rural Economic and Enterprise Development”, Working paper edition 1.0. Deutsche Gesellschaft für Technische Zusammenarbeit (GTZ), Eschborn, Germany: 138 pp.
- Hedges, L., Tipton, E., & Johnson, M.C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods*.
- Hombrados, J. G., and Waddington, H., (2012). *Internal Validity in Social Experiments and Quasi-Experiments: An Assessment Tool for Reviewers*. Mimeo: 3ie.
- International Labour Organization. (n.d.). *Working Conditions*. Retrieved from <http://www.ilo.org/global/topics/working-conditions/lang--en/index.htm>
- International Labour Organization. (2005). *Youth: Pathways to decent work*. ILC 93rd Session, Report VI Geneva.
- International Labour Organization. (2009). *Gender equality at the heart of decent work*. ILC 98th Session, Report VI Geneva.
- International Labour Organization. (2015). *Thesaurus*. Retrieved from <http://www.ilo.org>
- International Labour Organization. (2016). *Women at work: Trends 2016*. Retrieved from <http://www.refworld.org/docid/56de89c24.html>
- King, K., & Palmer, R. (2010). *Planning for technical and vocational skills development*. Paris, France: UNESCO International Institute for Educational Planning.

- Lipsey, M. W. & Wilson, D. B. (2001). *Practical Meta-Analysis*. Thousand Oaks, CA: SAGE.
- Ministerio de Trabajo, Colombia. (2016). *La Ley Projovent, Ley No. 1780*. <http://es.presidencia.gov.co/normativa/normativa/LEY%201780%20DEL%2002%20DE%20MAYO%20DE%202016.pdf>
- Patel, P. (2014). Effectiveness of entrepreneurship development interventions for women entrepreneurs: An ILO-WED Issue Brief. Retrieved from http://www.ilo.org/wcmsp5/groups/public/---ed_emp/---emp_ent/---ifp_seed/documents/publication/wcms_329356.pdf
- Piza, C., Cravo, T.A., Taylor, L., Gonzalez, L., Musse, I., Furtado, I., Sierra, A.C., Abdelnour, S. (2016). The impacts of business support services for small and medium enterprises on firm performance in low- and middle-income countries: A systematic review. *Campbell Systematic Reviews* 12(3).
- Pritchett, L. & Sandefur, J. (2013). *Context Matters for Size: Why External Validity Claims and Development Practice Don't Mix*. Center for Global Development, Working Paper 336.
- Revengea, A., M. Riboud, and H. Tan. (1994) *The Impact of Mexico's Retraining Program on Employment and Wages*. World Bank Economic Review, 8: 247-277. http://web.worldbank.org/archive/website00955A/WEB/PDF/MEXICO_T.PDF
- Revengea, A., & Shetty, S. (2012). Empowering women is smart economics—Closing gender gaps benefits countries as a whole, not just women and girls. *Finance and Development-English Edition*, 49(1), 40.
- Sandelowski, M., S. Docherty, & Emden, C. (1997). Focus on qualitative methods—Qualitative metasynthesis: Issues and techniques. *Research in Nursing and Health*, 20, 365–372.
- Shadish, W.R. (2011). Randomized controlled studies & alternative designs in outcome studies: Challenges and opportunities. *Research on Social Work Practice*, 21(6), 636-43.
- Snilstveit, B., Oliver, S., & Vojtkova, M. (2012). Narrative approaches to systematic review and synthesis of evidence for international development policy and practice. *Journal of Development Effectiveness*, 4(3), 409–429. Retrieved from <http://www.tandfonline.com/doi/pdf/10.1080/19439342.2012.710641>
- Snilstveit, B., Stevenson, J., Phillips, D., Vojtkova, M., Gallagher, E., Schmidt, T., Jobse, H., Geelen, M., Pastorello, M., and Eyers, J. (2015). *Interventions for improving learning outcomes and access to education in low - and middle - income countries: a systematic review*. 3ie Final Review. London: International Initiative for Impact Evaluation (3ie)
- Steiner, P.M., Cook, T.D., Shadish, W.R., & Clark, M.H. (2010). The importance of covariate selection in controlling for selection bias in observational studies *Psychological Methods* 15(3): 250-67.
- Sustainable Development Knowledge Platform (n.d.). *Sustainable development goals*. Retrieved from <https://sustainabledevelopment.un.org/sdgs>

- Trivedi, J.H. (1989). "Women's development through Distance Education" in ICDE Bulletin, Vol. 21, 1989, 17-22.
- UNCTAD (2016). Trade as a tool for the economic empowerment of women. Retrieved from http://unctad.org/meetings/en/SessionalDocuments/ciem8d2_en.pdf
- Waddington, H., Snilstveit, B., Hombrados, J. G., Vojtkova, M., Anderson, J., & White, H. (2014). *Farmer field schools for improving farming practices and farmer outcomes in low- and middle-income countries: A systematic review*. Oslo, Norway: The Campbell Collaboration. Retrieved from <http://www.campbellcollaboration.org/lib/project/203/>
- Waddington, H., White, H., Snilstveit, B., Hombrados, J., Vojtkova, M., Davies, et al. (2012). How to do a good systematic review of effects in international development: A tool kit. *Journal of Development Effectiveness*, 4(3), 359–387. Retrieved from <http://dx.doi.org/10.1080/19439342.2012.71176>
- Walsh, D. & Downe, S. (2005). Meta-synthesis method for qualitative research: A literature review. *Journal of Advanced Nursing*, 50(2), 204–211.
- White, H. (2009). Theory-based impact evaluation: Principles and practice. *Journal of Development Effectiveness*, 1(3), 271–284.
- World Bank (n.d.). Jordan: New Work Opportunities for Women Pilot Project. Retrieved from <http://web.worldbank.org/WBSITE/EXTERNAL/COUNTRIES/MENAEXT/0,,contentMDK:22872667~menuPK:3949143~pagePK:146736~piPK:226340~theSitePK:256299,00.html>
- Zimmer, L. (2006). Qualitative meta-synthesis: A question of dialoguing with texts. *Journal of Advanced Nursing*, 53(3), 311–318.

REVIEW AUTHORS

Lead review author:

The lead author is the person who develops and coordinates the review team, discusses and assigns roles for individual members of the review team, liaises with the editorial base, and takes responsibility for the ongoing updates of the review.

Name:	Marjorie Chinen
Title:	Principal Researcher
Affiliation:	American Institutes for Research
Address:	1000 Thomas Jefferson Street, NW
City, State, Province or County:	Washington, DC
Postal Code:	20007
Country:	USA
Phone:	(202) 403-5171

Email: mchinen@air.org

Co-author(s):

Name: Thomas de Hoop

Title: Senior Researcher

Affiliation: American Institutes for Research

Address: 1000 Thomas Jefferson Street, NW

City, State, Province or County: Washington, DC

Postal Code: 20007

Country: USA

Phone: (202) 403-5803

Email: tdehoop@air.org

Name: María Balarin

Title: Principal Researcher

Affiliation: Group for the Analysis of Development (GRADE)

Address: Av. Almirante Grau 915

City, State, Province or County: Lima

Postal Code: 15063

Country: Peru

Phone: (511) 247-9988

Email: mbalarin@grade.org.pe

Name: Lorena Alcázar

Title: Principal Researcher

Affiliation: Group for the Analysis of Development (GRADE)

Address: Av. Almirante Grau 915

City, State, Province or County: Lima

Postal Code: 15063

Country: Peru

Phone: (511) 247-9988

Email: lalcazar@grade.org.pe

ROLES AND RESPONSIBILITIES

- Content: María Balarin, Lorena Alcázar
- Systematic review methods: Marjorie Chinen, Thomas de Hoop
- Statistical analysis: Thomas de Hoop, Marjorie Chinen
- Information retrieval: John Eyers

APPENDIX 1: EXAMPLE SEARCH STRATEGY (ECONLIT)

Search	Query
S20	S3 and S6 and S9 and S19 – Final result
S19	or/S10-S18
S18	developing countries.hw.
S17	transitional countr*.ti,ab.
S16	(Imic or Imics or third world or lami countr*).ti,ab.
S15	(low adj3 middle adj3 countr*).ti,ab.
S14	(low* adj (gdp or gnp or gross domestic or gross national)).ti,ab.
S13	((developing or less* developed or under developed or underdeveloped or middle income or low* income) adj (economy or economies)).ti,ab.
S12	((developing or less* developed or under developed or underdeveloped or middle income or low* income or underserved or under served or deprived or poor*) adj (countr* or nation? or population? or world or state*)).ti,ab.
S11	(Afghanistan or Albania or Algeria or Angola or Argentina or Armenia or Armenian or Azerbaijan or Bangladesh or Benin or Byelarus or Byelorussian or Belarus or Belorussian or Belorussia or Belize or Bhutan or Bolivia or Bosnia or Herzegovina or Hercegovina or Botswana or Brazil or Bulgaria or Burkina Faso or Burkina Fasso or Upper Volta or Burundi or Urundi or Cambodia or Khmer Republic or Kampuchea or Cameroon or Cameroons or Cameroon or Camerons or Cape Verde or Central African Republic or Chad or China or Colombia or Comoros or Comoro Islands or Comores or Mayotte or Congo or Zaire or Costa Rica or Cote d'Ivoire or Ivory Coast or Croatia or Cuba or Djibouti or French Somaliland or Dominica or Dominican Republic or East Timor or East Timur or Timor Leste or Ecuador or Egypt or United Arab Republic or El Salvador or Eritrea or Ethiopia or Fiji or Gabon or Gabonese Republic or Gambia or Gaza or Georgia Republic or Georgian Republic or Ghana or Gold Coast or Grenada or Guatemala or Guinea or Guiana or Guyana or Haiti or Honduras or Hungary or India or Maldives or Indonesia or Iran or Iraq or Jamaica or Jordan or Kazakhstan or Kazakh or Kenya or Kiribati or Korea or Kosovo or Kyrgyzstan or Kirghizia or Kyrgyz Republic or Kirghiz or Kirgizstan or Lao PDR or Laos or Lebanon or Lesotho or Basutoland or Liberia or Libya or Macedonia or Madagascar or Malagasy Republic or Malaysia or Malaya or Malay or Sabah or Sarawak or Malawi or Mali or Marshall Islands or Mauritania or Mauritius or Agalega Islands or Mexico or Micronesia or Middle East or Moldova or Moldavia or Moldovian or Mongolia or Montenegro or Morocco or Ifni or Mozambique or Myanmar or Myanma or Burma or Namibia or Nepal or Netherlands Antilles or New Caledonia or Nicaragua or Niger or Nigeria or Pakistan or Palau or Palestine or Panama or Paraguay or Peru or Philippines or Philipines or Phillipines or Phillippines or Papua New Guinea or Romania or Rumania or Roumania or Rwanda or Ruanda or Saint Lucia or St Lucia or Saint Vincent or St Vincent or Grenadines or Samoa or Samoan Islands or Navigator Island or Navigator Islands or Sao Tome or Senegal or Serbia or Montenegro or Seychelles or Sierra Leone or Sri Lanka or Solomon Islands or Somalia or Sudan or Suriname or Surinam or Swaziland or South Africa or Syria or Tajikistan or Tadzhiestan or Tadjikistan or Tadjik or Tanzania or Thailand or Togo or Togoese Republic or Tonga or Tunisia or Turkey or Turkmenistan or Turkmen or Uganda or Ukraine or Uzbekistan or Uzbek or Vanuatu or New Hebrides or Venezuela or Vietnam or Viet Nam or West Bank or Yemen or Yugoslavia or Zambia or Zimbabwe).ti,ab,hw,ct.
S10	(Africa or Asia or Caribbean or West Indies or South America or Latin America or Central America).ti,ab,hw,gr.
S9	S7 or S8
S8	(J24 or I24 or L26 or I23 or I26 or I21 or P36 or M53).cc.

Search	Query
S7	(skill* or upskill* or "life skills" or life-skills or train* or retrain* or mentor* or apprentice* or intern or interns or internship* or value* or prestige or expertise or professional* or qualifi* or educat* or coaching or leadership or entrepreneur* or "human capital" or (business* adj2 (manag* or support*)) or (capacity adj2 (build* or acqui*)) or projoven).ti,ab.
S6	S4 or S5
S5	(J31 or J71 or J44 or G21 or J78).cc.
S4	(employment or employab* or employer* or employee* or work or workforce or job* or vocation* or career* or occupation* or livelihood* or workplace* or part-time or casual or informal* or wages or ((labour or labour) adj market*) or school-to-work or "school to work").ti,ab.
S3	S1 or S2
S2	J16.cc.
S1	(wom?n or female* or girl* or schoolgirl* or wife or wives or mother* or gender* or (occupation* adj2 (segregation or stratification)) or empower* or businesswomen or businesswoman or (wom?n adj2 (business* or industr* or enterpris* or commerc*))).ti,ab.

APPENDIX 2: PILOT PHASE 1 SCREENING

If any of the preliminary exclusion criteria can be checked, the study is excluded and screening is stopped. For each of the criteria, do not exclude a study if there is not enough information in the title and abstract to exclude with certainty.

1. **Exclude** if date of publication is before 1990.
2. **Exclude** if language is not English or Spanish.
3. **Exclude** if setting is a high-income country.
4. **Exclude** if there is no reference to women, gender, etc.

Studies that remain are then assessed for their relevance to the primary or secondary research questions. Studies may be relevant to more than one of the research questions, for example, if they have a mixed-methods design or explore the secondary research question.

If a study is quantitative primary evidence, use the exclusion criteria that immediately follow. After the study is either included or excluded based on these criteria, proceed to screen the study on the basis of the second research question. If a study passes any of the three sets of inclusion criteria, it will be included in Phase 2 for full text screening.

5. Quantitative primary study (research question 1)
 - 5a. **Exclude** if the study does not focus on evaluating an intervention.
 - 5b. **Exclude** if the study does not use longitudinal or cross-sectional data collected at the individual level.
 - 5c. **Exclude** if the study design is not experimental or if the study is quasi-experimental without propensity score or another type of matching, difference-in-difference estimation, instrumental variable regression, multivariate cross-sectional regression analysis, or other forms of

multivariate analysis (such as the Heckman selection model or multivariate ordinary least squares regression analysis).

- 5d. **Exclude** if the study does not include a valid comparison condition, such as “no intervention,” “pipeline,” or “business as usual.”
- 5e. **Exclude** if the study does not include any of the following outcome measures.
- 5f. **Include** for research question 1 if the study passes all of the previous criteria.

6. Qualitative primary study (research question 2)

- 6a. **Exclude** if the study is not tied to specific vocational or business training interventions.
- 6b. **Exclude** if the study does not report a qualitative methodology, such as interviews or focus groups.
- 6c. **Exclude** if the study does not include women.
- 6d. **Include** for research question 2 if the study passes all of the previous criteria.

7. **Include** for Phase 2 if any of 5f, 6d, or 7d are checked.

APPENDIX 3: DATA EXTRACTION FORM FOR STUDIES INCLUDED TO ADDRESS THE PRIMARY RESEARCH QUESTION

Study Data Extraction/Coding

Study ID (sid):

Coder's initials (coderid):

Date coded (date):

Author(s):

General Comments:

Publication date:

Publication type: (1) book, (2) peer-reviewed journal, (3) book chapter, (4) dissertation/thesis, (5) unpublished report

Program Data Extraction/Coding

Intervention level:

Intervention Components (more than 1 answer possible): (1) courses in administrative occupations (e.g. marketing, secretarial work, sales), (2) courses in manual occupations (e.g. electrician, cooking assistant), (3) courses in fairly skilled occupations (e.g. account assistant, information technology specialist), (4) internships, (5) on-the-job training, (6) connecting participants with potential employers, (7) development of job-related life skills (e.g. CV training, interpersonal relationships), (8), development of broader life skills (e.g. reproductive

health, household economics), (9) general business training (e.g. keeping business records and encouraging small business owners to separate household and business finances, (10) technical business training, (11) sector-specific business training, (12) changing entrepreneurial attitudes or aspirations, (13) knowledge about the value chain of the product, (14) asset transfers, (15) in-kind transfers, (16) cash transfers, (17) saving accounts, (18) asset-specific training, (19) regular follow-up visits by assets specialists, (20) stipends for women to participate in training, (21) additional support for women with children, (22) negotiation support with families to ensure participation in training, (23 other: define)

Name of intervention:

Implementer:

Intervention Description:

Scale of the intervention:

Intervention frequency/duration/intensity:

Intervention start date:

Intervention end date:

Is there information on program take-up or adherence?:

Which methods are used to assess program take-up or adherence?:

Are there any unit cost data/cost-effectiveness estimates provided?:

If so, report any details on unit cost and/or total cost:

Countries:

World Bank Region(s):

Is the study conducted in rural areas?:

Is the study conducted in urban areas?:

Percent of the sample female:

Average age of sample participants:

Primary outcomes measured:

Author's definition of primary outcomes:

Secondary outcomes measured:

Author's definition of secondary outcomes:

Data collection method:

Frequency of outcome data collection:

Level of assignment to treatment and control group:

Implementer of the program:

Sampling frame for selection of study participants:

How was the population sampled and why?:

Target group (men, women or both):

Study design:

Sample size:

Effect Size Extraction/Coding

Outcome category (outcat): (1) economic (2) political (3) social (4) psychological

Outcome name (outname):

Direction of effect (esdir): (1) effect favors self-help group (2) effect favors comparison (3) effect favors neither (4) cannot tell

Effect is statistically significant (essig)?: (1) yes (2) no (3) cannot tell
Treatment sample size (shgss):
Comparison sample size (compss):

For continuous measures:

Treatment group mean (txmean):
Comparison group mean (compmean):
Are means reported above adjusted? (meanadj): (1) yes (2) no

Treatment group standard deviation (txsd):
Comparison group standard deviation (compsd):

Treatment group standard error (txse):
Comparison group standard error (compse):

t-value from an independent *t*-test (est)

For regression estimates:

Regression coefficient: (regress):
Regression standard error: (regresses):

For dichotomous measures:

Treatment group number of participants who experienced a change (txnum):
Comparison group number of participants who experienced a change (compnum):

Treatment group proportion of participants who experienced a change (txpro):
Comparison group proportion of participants who experienced a change (comppro):

Are the proportions above adjusted for pretest variables? (proadj): (1) yes (2) no

Logged odds-ratio (eslgodd):

Standard error of logged odds-ratio (eslgoddse):

Logged odds-ratio adjusted? (e.g., from a logistic regression analysis with other independent variables) (1=yes; 0=no)

Chi-square value with $df = 1$ (2 by 2 contingency table) (eschi):
Correlation coefficient (esphi):

APPENDIX 4: DATA EXTRACTION FORM FOR STUDIES INCLUDED TO ADDRESS THE SECONDARY RESEARCH QUESTIONS

Study Data Extraction/Coding

Study ID (sid):

Coder's initials (coderid):

Date coded (date):

Author(s) (author):

Publication date (pubdate):

Country (country):

Start date of study (startdate):

End date of study (enddate):

Publication type (pubtype): (1) book, (2) peer-reviewed journal, (3) book chapter, (4) dissertation/thesis, (5) unpublished report

Document type (doctype): (1) qualitative study, (2) process evaluation, (3) project document

Program Data Extraction/Coding

Study ID (sid):

Coder's initials (coderid):

Name of Study Evaluated under Research Question 1:

Date coded (date):

Name of vocational or business training program:

Vocational or business training:

Intervention Components (more than 1 answer possible): (1) courses in administrative occupations (e.g. marketing, secretarial work, sales), (2) courses in manual occupations (e.g. electrician, cooking assistant), (3) courses in fairly skilled occupations (e.g. account assistant, information technology specialist), (4) internships, (5) on-the-job training, (6) connecting participants with potential employers, (7) development of job-related life skills (e.g. CV training, interpersonal relationships), (8), development of broader life skills (e.g. reproductive health, household economics), (9) general business training (e.g. keeping business records and encouraging small business owners to separate household and business finances, (10) technical business training, (11) sector-specific business training, (12) changing entrepreneurial attitudes or aspirations, (13) knowledge about the value chain of the product, (14) asset transfers, (15) in-kind transfers, (16) cash transfers, (17) saving accounts, (18) asset-specific training, (19) regular follow-up visits by assets specialists, (20) stipends for women to participate in training, (21) additional support for women with children, (22) negotiation support with families to ensure participation in training, (23 other: define)

Organization that implements vocational or business training program:

Descriptive details about what is delivered to participants:

Descriptive details about who delivers the Intervention (profession training level, number of staff etc.)

Descriptive details about duration/frequency/intensity of training:

Start date of intervention:

End date of intervention:

Sector:

Target Group: (1) Women only, (2) Women and Men

Country:

Rural or Urban Area: (1) Rural, (2) Urban, (3) Both Rural and Urban

Age of beneficiaries:

Describe the Costs of the Intervention:

For Qualitative Data:

Data Collection Type: (1) Focus Groups, (2) In-Depth Interviews, (3) Ethnographies, (4) Participatory Research, (5) Other: Define

Sampling: (1) Random, (2) Purposive

Sample Size:

Description of Sampling Procedure:

Description of Themes:

Description of Analysis:

APPENDIX 5: RISK BIAS TOOL CATEGORIES FOR QUANTITATIVE STUDIES

Risk of Bias Assessment Tool	
Justification of use	
Which primary outcomes are measured in the study?	
Provide the authors definition of each included primary outcome	
Which secondary outcomes are measured in the study?	
Provide the authors definition of each included secondary outcome	
Describe methods of data collection	
What is the frequency of outcome data collection?	
At which level was assignment to treatment and control group conducted?	
Does the study show baseline values of the outcomes of interest (as defined in the protocol) for beneficiaries and non-beneficiaries?	
If baseline values of the outcome of interest are not available at baseline, does the study show baseline values of characteristics of beneficiaries and non-beneficiaries that are not likely to be affected by the intervention?	
Are the mean values or the distributions of the covariates at baseline statistically different for beneficiaries and non-beneficiaries ($p < 0.05$)?	
If there are statistically significant differences between beneficiaries and non-beneficiaries are these differences controlled for using covariate analysis in the impact evaluation?	
If baseline characteristics are not available, does the study qualitatively assess why beneficiaries are likely/unlikely to be a random draw of the population at baseline?	
Confounding and selection bias	
Does the study use a comparison/control group of women without access to the program?	
Does the study use a comparison/control group of women with access to the program but that did not choose to participate in the program?	
Does the study include data on the outcomes of interest at baseline and endline (before and after the intervention)?	
Are the data on covariates collected at the baseline?	

Is difference-in-difference estimation used?	
If the study is quasi-experimental and uses difference-in-difference estimation do the authors assess the parallel trends assumption?	
If the study does not use difference-in-difference, does the study control for baseline values of the outcome of interest (ANCOVA)	
If the study does not use difference-in-difference and does not control for baseline values of the outcome variable, does the study control for other covariates at baseline	
If the study does not use difference in differences estimation, is there any assessment of likely risk of bias from time invariant characteristics driving both participation and outcome?	
If the study does not use difference in difference estimation but does assess likely risk of bias from time invariant characteristics, are these time invariant characteristics likely to bias the impact estimates	
Does the study report the table with the results of the outcome equation (including covariates)?	
Where full results of the outcome equation are not reported, is it clear which covariates have been used?	
Are all relevant observable covariates (confounding variables) included in the outcome equation which might explain outcomes, if estimation does not use a statistical technique to control for selection bias (RCT, PSM, RDD, or IV)?	
Attrition	
For studies including baseline data, does the study report attrition (drop-out) from the study?	
Is the attrition rate from the study below 10%?	
Does the study assess whether drop-outs from the study are random draws from the sample (e.g. by examining correlation with determinants of outcomes, in both treatment comparison group)?	
Spillovers and contamination	
Spillovers: are comparisons sufficiently isolated from the intervention (eg participants and non-participants are sufficiently geographically or socially separated) or are spillovers estimated by comparing non-beneficiaries with access to the intervention to non-beneficiaries without access to the intervention and/or through social network analysis?	
Spillovers; if spillovers are not estimated, is the study likely to overestimate or underestimate the impact of the program?	
Contamination: does the study assess whether the control group receives the intervention?	
Contamination: if the control group receives the intervention but for a shorter amount of time does the study assess the likelihood that the control group has received equal benefits as the treatment group	
Contamination: if the control group receives the intervention have they received the intervention sufficiently long to argue that they have benefited from the intervention	
Other threats to validity	

Does the evidence suggest analysis reporting biases are likely/unlikely to be serious? Analysis reporting biases include failure to report important treatment effects (possibly relating to intermediate outcomes), or justification for (uncommon) estimation methods, especially multivariate analysis for outcomes equations.	
Hawthorne and John Hendry Effects	
Do the authors argue convincingly that it is not likely that being monitored influences the behaviour of the beneficiaries and non-beneficiaries in different ways?	
Confidence Intervals	
Does the study account for lack of independence between observations within assignment clusters if the outcome variables are clustered?	
Is the sample size likely to be sufficient to find significant effects of the intervention?	
Do the authors control for heteroskedasticity and/or use robust standard errors?	
Ask questions below only for studies that apply randomization	
Does the study apply randomized assignment?	
Does the study use a unit of allocation with a sufficiently large sample size to ensure equivalence between the treatment and the control group	
Ask questions below only for studies that apply regression discontinuity designs	
Is the allocation of the program based on a pre-determined continuity on a continuous variable and blinded to the beneficiaries or if not blinded, individuals cannot reasonably affect the assignment variable in response to knowledge of the participation rule?	
Is the sample size immediately at both sides of the cut-off point sufficiently large to equate groups on average?	
Is the mean of the covariates of individuals immediately at both sides of the cut-off point statistically significantly different for beneficiaries and non-beneficiaries?	
If there are statistically significant differences between beneficiaries and non-beneficiaries are these differences controlled for using covariate analysis?	
Ask questions below only for studies that apply matching	
Quality of matching (PSM, covariate matching)	
Are beneficiaries and non-beneficiaries matched on all relevant characteristics?	
Does the study report the results of the matching function (eg for PSM the logit function)?	
Does the study report the matching method?	
Does the study exclude observations outside the common support?	
Does the study use variables at follow-up that can be affected by the intervention in the matching equation?	
Does the study report the mean or distribution for the covariates of the treatment and control groups after matching?	

Are these characteristics similar, based on tests for statistically significant differences ($p > 0.05$)?	
Ask questions below only for studies that apply instrumental variable estimation	
Does the study describe clearly the instrumental variable(s)/identifier used?	
Are the results of the participation equation reported?	
Are the instruments jointly significant at the level of $F \geq 10$? If an F test is not reported, does the author report and assess whether the R-squared of the instrumenting equation is large enough for appropriate identification ($R\text{-sq} > 0.5$)?	
Are the instruments individually significant ($p \leq 0.05$)?	
For IV, If more than one instrument is used in the procedure, does the study include and report an overidentifying test ($p \leq 0.05$ is required to reject the null hypothesis)?	
Does the study qualitatively assess the exogeneity of the instrument/identifier (both externality as well as why the variable should not enter by itself in the outcome equation)?	
Ask questions below only for studies with censored outcome variables	
Do the authors use appropriate methods (e.g. Heckman selection models, tobit models, duration models) to account for the censoring of the data?	
For Heckman models; is there is a variable that is statistically significant in the first stage of the selection equation and excluded from the second stage	
Risk of Selection Bias	High/Medium/Low
Risk of Performance Bias	High/Medium/Low
Risk of Outcome and Analysis Reporting Biases	High/Medium/Low
Risk of Other Biases	High/Medium/Low

SOURCES OF SUPPORT

This review is supported by external funding from the International Initiative for Impact Evaluation (3ie).

DECLARATIONS OF INTEREST

None of the proposed authors have developed studies that focus on the topic area. However, María Balarin and Lorena Alcázar are part of GRADE, an institution that has published several studies on the effectiveness of vocational and business training in the context of Latin America. We will mitigate this concern by placing a greater emphasis on the assessment of these studies by the non-GRADE authors.

PRELIMINARY TIMEFRAME

Deliverable	Date to Submit
Title Registration	January 2016
Protocol	March 2016
Draft review	November 2016
Final review	February 2017

PLANS FOR UPDATING THE REVIEW

The authors will assess updating the review every three years if funding becomes available.

ACKNOWLEDGMENTS

Our sincere appreciation goes to the International Initiative for Impact Evaluation (3ie), Department of Foreign Affairs and Trade (DFAT) Canada, and UN Women for the financial and technical support provided. Special thanks go to John Eyers for his expert guidance and support on searching the literature. In addition, we would like to thank Joshua Sennett, from AIR, for his assistance on various aspects of this review.

AUTHOR DECLARATION

Authors' responsibilities

By completing this form, you accept responsibility for preparing, maintaining and updating the review in accordance with Campbell Collaboration policy. The Campbell Collaboration will provide as much support as possible to assist with the preparation of the review.

A draft review must be submitted to the relevant Coordinating Group within two years of protocol publication. If drafts are not submitted before the agreed deadlines, or if we are unable to contact you for an extended period, the relevant Coordinating Group has the right to de-register the title or transfer the title to alternative authors. The Coordinating Group also has the right to de-register or transfer the title if it does not meet the standards of the Coordinating Group and/or the Campbell Collaboration.

You accept responsibility for maintaining the review in light of new evidence, comments and criticisms, and other developments, and updating the review at least once every five years, or, if requested, transferring responsibility for maintaining the review to others as agreed with the Coordinating Group.

Publication in the Campbell Library

The support of the Coordinating Group in preparing your review is conditional upon your agreement to publish the protocol, finished review, and subsequent updates in the Campbell Library. The Campbell Collaboration places no restrictions on publication of the findings of a Campbell systematic review in a more abbreviated form as a journal article either before or after the publication of the monograph version in *Campbell Systematic Reviews*. Some journals, however, have restrictions that preclude publication of findings that have been, or will be, reported elsewhere and authors considering publication in such a journal should be aware of possible conflict with publication of the monograph version in *Campbell Systematic Reviews*. Publication in a journal after publication or in press status in *Campbell Systematic Reviews* should acknowledge the Campbell version and include a citation to it. Note that systematic reviews published in *Campbell Systematic Reviews* and co-registered with the Cochrane Collaboration may have additional requirements or restrictions for co-publication. Review authors accept responsibility for meeting any co-publication requirements.

I understand the commitment required to undertake a Campbell review, and agree to publish in the Campbell Library. Signed on behalf of the authors:

Form completed by:

**Date: 30 September
2016**