International Initiative for Impact Evaluation
Working Paper 3

المبادرة الدولية لتقييم الأثر

ورقة عمل 3

تقييم الأثر المرتكز على النظرية: مبادئ وممارسات

هوارد وايت ، يونيو 2009

مبادرة 3ie :

المبادرة الدولية لتقييم الأثر (المردود) منظمة دولية غير حكومية توفر الدعم المادي لتشجيع السياسات والبرامج التنموية. وهى تعتبر المنظمة الدولية الرائدة في توفير الدعم وإنتاج الأدلة ذات الجودة العالية على : ماذا يعمل وكيف يعمل ولما يعمل وبأي تكلفة بالنسبة الى أثر ومردود التنمية. ونحن نعتقد أن الأدلة الأفضل والمبينة على السياسات الجيدة سوف تجعل التنمية أكثر فعالية، وتحسن من حياة المجتمعات.

أوراق عمل المبادرة :

تركز هذه الأوراق على القضايا الحالية والمناقشات والتحديات التي تواجه مُعدي السياسات والعاملون في مجال تقييم الأثر، والقائمون على المراجعة المنهجية في هذا المجال. تعتمد أوراق العمل المرتبطة بالسياسات على النتائج ذات الصلة الناتجة عن عمليات تقييم الأثر والمراجعات المنهجية الممولة من المبادرة، ونتائج عن تقييمات دقيقة موثوق بها، ومراجعات تُقدم رؤى، أو تحليل جديد، أو نتائج وتوصيات. وتعتمد الأوراق التي تُركز على الطرق على مصادر مشابهة مساعدة للفهم، والتصميم، والإستخدام المتقدم للتقييم، والمراجعة الدقيقة والملائمة.

حول هذه الورقة :

نشرت هذه الورقة لأول مرة باللغة الإنجليزية في عام 2009. ولم يحدث اي تعديل أو إضافات على الإصدار الأصلي قبل ترجمته الى اللغة العربية. وقد قام بترجمة هذه الورقة الدكتور / محمد مصطفى الفولي والسيد / تامر محمد الفولي، وتقوم المبادرة بإتاحة هذه الورقة بدون اي مراجعة للترجمة، وأن المحتوى بالكامل، واي إضافات أو حذف هى مسئولية المؤلفين ولا تعبر عن رأي المبادرة، أو المانحين، أو مجلس المفوضين. وأي أخطاء أو حذف حدث خلال الترجمة – تعد مسئولية المترجمين.

يمكن الإشارة الى هذه الورقة كما يلي :

Suggested citation : White, H. 2009. *Theory-based impact evaluation, principles and practice, 3ie Working paper 3.* New Delhi, International Intitative for Impact Evaluation (3ie).

المحررين الحصرين لسلسلة أوراق العمل : بيرلي ليتش - عمانويل جيمينيز

مدير التحرير: كانيكا جها

المترجمون : د. محمد مصطفى الفولي / تامر محمد الفولي

الصفحة الخارجية : البرنامج الإنمائي للأمم المتحدة

© International Initiative for Impact evaluation (3ie), 2015

تقييم الأثر المرتكز على النظرية: مبادئ وممارسات

هوارد وايت

المدير التنفيذي للمبادرة الدولية لتقييم الأثر

البريد الإلكتروني: hwhite@3ieimpact.org

ملخص

اقترنت دعوات تقييم الأثر ليس فقط لمعرفة ما يفيد (يعمل)، ولكن ايضا لماذا يفيد (يعمل). إن النهج المرتكز على النظرية لتقييم الأثر والمقبول على نطاق واسع سيسلط الضوء على السؤال لماذا. فالنظرية توضح السلسلة السببية من المدخلات إلى التأثير والآثار، وتختبر الافتراضات الأساسية. لكن تطبيق النهج المرتكز على النظرية لا يزال ضعيفا. هذه الورقة تحدد المبادئ الستة التالية للتطبيق الناجح للنهج: (1) توضيح السلسلة السببية (نظرية البرنامج)، (2) فهم السياق؛ (3) توقع عدم التجانس، (4) تقييم دقيق للأثر باستخدام الواقع المضاد الموثوق به، (5) تحليل واقعي دقيق، و(6) استخدام وسائل مختلطة

1- مقدمة

شهدت السنوات الأخيرة اهتماما متزايدا في استخدام اساليب التحليل الكمي لقياس أثر برامج التنمية، مثل برامج عمل المنظمات (مثال ذلك: مختبر مكافحة الفقر[1]، والابتكارات في مجال مكافحة الفقر[2])، ومجموعة من الدراسات الممولة تحت مظلة مبادرة البنك الدولي لتقييم أثر التنمية[3] والصندوق الاسباني لتقييم الأثر[4]، والتمويل المتاح من قبل المبادرة الدولية لتقييم الأثر[5]. كل هذا يعني أنه خلال الخمس سنوات القادمة سيكون هناك المئات من هذه الدراسات، مقارنة بالحفنة المذكورة في الدراسات التي أجريت في السنوات الأخيرة (مثل: مركز التنمية العالمية[6]، 2006). ومع ذلك، فإن شعار معظم من يدعمون التحرك نحو تقييم أفضل للأثر هو "فهم ليس فقط ما يفيد (يعمل)، ولكن لماذا يفيد (يعمل)". هذه الرؤية لا تمنح من خلال تقديم تقارير بسيطة عن متوسط اثر التدخل. ومن هنا جاء بيان شبكة شبكات تقييم الأثر[7]: "تطبيق النهج المرتكز على النظرية يعني أن التصميم الجيد لتقييم الأثر يغطي كل من العملية وأسئلة تقييم الأثر. مما يعزز بالتالي الصلة بالسياسات حيث أن الدراسات يمكنها معالجة أسئلة: لماذا - أو لماذا لا - يحقق التدخل الأثر المنشود، وليس ما إذا كان قد تحقق" (NONIE، بدون تاريخ). وبالمثل، نجد أن دليل المبادرة لممارسات تقييم الأثر تقول أن "الدراسات يجب أن تبين بوضوح كيف من المتوقع ان التدخل (المدخلات) سيؤثر على التأثير النهائي، واختبار كل ارتباط (فرضية) من المدخلات إلى التأثير (التي يشار إليها أحيانا بنظرية البرنامج) . وينبغي أن يُدرج تصميم التقييم تحليل السلسلة السببية من المدخلات إلى الآثار" (3ie، بدون تاريخ: 2).

النهج الذي ننادى به هنا "أن نفهم لماذا تحقق / لم يتحقق الأثر" المقصود به هو تقييم الأثر المرتكز على النظرية[8]. ليس هناك شيء جديد حول هذا الموضوع. التقييم المرتكز على النظرية - وهو يعني النظر في الفرضيات التي تقوم عليها السلسلة السببية من المدخلات إلى التأثير والأثر - هو نهج راسخ (انظر على سبيل

[1] مختبر مكافحة الفقر www.povertyactionlab.org – Poverty Action Lab (J-PAL)
[2] الابتكارات في مجال مكافحة الفقر http://poverty-action.org – Innovations in Poverty Action (IPA)
[3] مبادرة البنك الدولي لتقييم أثر التنمية – World Bank's Development Impact Evaluation Initiative (DIME)
[4] صندوق الاسباني لتقييم الأثر– Spanish Impact Evaluation Fund (SIFE)
[5] المبادرة الدولية لتقييم الأثر – www.3ieimpact.org – International Initiativefor Impact Evaluation (3ie)
[6] مركز التنمية العالمية – Centre for Global Development
[7] شبكة شبكات تقييم الأثر – Network of Networks on Impact Evaluation (NONIE)
[8] تقييم الأثر المرتكز على النظرية – theory-based impact evaluation (TBIE)

المثال، فايس عام 1998، وكارفاليو ووايت، 2004، للتطبيق في إطار أنشطة تنموية). تستخدم توضيحات نظرية البرنامج منذ فترة طويلة من قبل بعض الممارسين للنهج التجريبي وشبه التجريبي، باعتباره وسيلة لشرح النتائج التي توصلوا إليها (بلاكمان ورايخ، 2009: 67-68). في ورقتها لمراجعة تصميم ممكن لتقييم الأثر لمجموعة من التدخلات التنموية، (روجرز، 2009) تلاحظ أن النهج المرتكز على النظرية سيكون من المناسب في كل حالة.

على الرغم من أن الالتزام بتقييم الأثر المرتكز على النظرية موجود من حيث المبدأ، إلا أن عدد قليل من الدراسات تظهر الوفاء بالوعد لهذا النهج في الممارسة. وتهدف هذه الورقة إلى المساعدة في تخطي هذه الفجوة من خلال وضع الخطوات أو المبادئ وراء التقييم المرتكز على النظرية. أبدأ في الجزء الثاني بمثال: مشروع التغذية المتكاملة ببنجلاديش[9]، والتي سيتم اتباعها بعد ذلك بأمثلة أخرى، لتوضيح المبادئ التي تناقش في الجزء الثالث. ويقارن الجزء الرابع باختصار تقييم الأثر المرتكز على النظرية بنهج الصندوق الاسود ويشمل الجزء الخامس الخلاصة.

2. مثال: مشروع التغذية المتكاملة ببنجلاديش

يقدم هذا القسم لمحة موجزة عن تقييم مشروع التغذية المتكاملة ببنجلاديش. وسيتم استخدام هذا المثال لتوضيح مبادئ تقييم الأثر المرتكز على النظرية التي سيتم مناقشتها في القسم التالي. ويمكن الاطلاع على مزيد من المناقشات الموسعة لهذا المشروع في البنك الدولي (2005)، وايت وماسيه (2006) ووايت (2005).

وكان مشروع التغذية المتكاملة ببنجلاديش مبنيا على غرار مشروع التغذية المتكاملة بتاميل نادو[10] المشهود له في الهند، وهو مشروع لرصد نمو الأطفال. وكان يتم وزن الأطفال الرضع أسبوعيا في محطة محلية تعمل بها امرأة قروية مدربة لتكون ممارس تغذية في المجتمعات المحلية. ويتم اسقاط الوزن امام العمر على مخطط

[9] مشروع التغذية المتكاملة ببنجلاديش – Bangladesh Integrated Nutrition Project (BINP)
[10] مشروع التغذية المتكاملة بتاميل نادو – Tamil Nadu Integrated NutritionProject (TINP)

نمو، ثم ادراج الأطفال الذين كان نموهم ضعيف (نمو متعثر) أو أقل من المعيار المرجعي بكثير (سوء التغذية) في البرنامج. يتألف البرنامج من نصائح (مشورة) غذائية وتغذية تكميلية. ومع ذلك، كانت وثائق المشروع توضح أن الأثر الرئيسي المتوقع سيكون من خلال تقديم النصائح (المشورة) الغذائية. وقد كان المبرر أن الجهل، وليس الفقر، هو المسؤول عن سوء التغذية، وهي حجة مدعومة ببيانات تظهر سوء التغذية حتى في الشرائح الأغنى، ووجود معتقدات مثل "أكل أقل"، وهو يعني أن المرأة يجب أن تأكل أقل خلال فترة الحمل. استهدف البرنامج أيضا النساء الحوامل مع المشورة الغذائية والتغذية التكميلية. وكان مشروع التغذية المتكاملة ببنجلاديش برنامجا تجريبيا ، وتبعه في وقت لاحق برنامج التغذية الوطني.

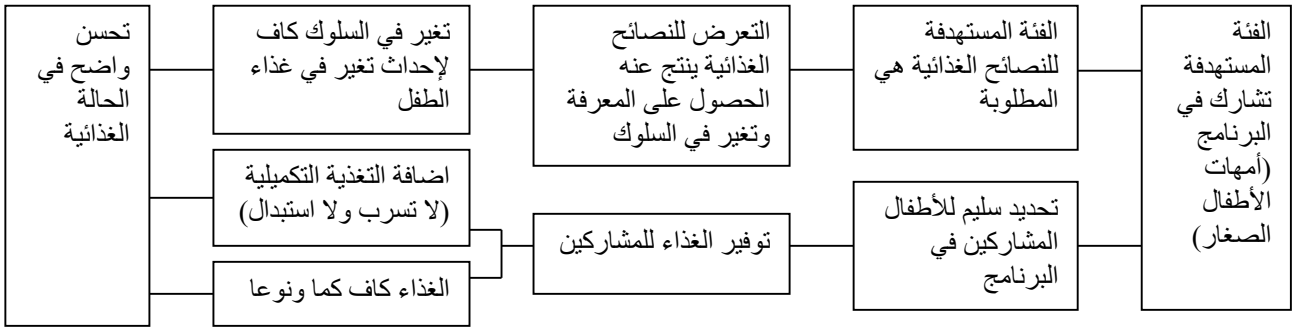
كان الاعتقاد في البداية أن مشروع التغذية المتكاملة ببنجلاديش ناجحا. وقد أظهرت بيانات المتابعة في مناطق عمل المشروع انخفاضات كبيرة في سوء التغذية، ولا سيما سوء التغذية الحاد. وفي حوالي منتصف مدة المشروع، وعلى أساس هذه الدلائل وقبل أي تقييم، قرر البنك أن يتوسع مع برنامج التغذية الوطني. أصدرت هيئة إنقاذ الطفولة - المملكة المتحدة تقريرا ينتقد هذا القرار مقدمة البيانات الخاصة بها مقارنة بين التصميم البسيط اللاحق للتدخل (بيانات مجموعة التدخل) مقابل تصميم المراجعة (مجموعة المقارنة) الذي لم يظهر أي فرق بين المنطقتين (هيئة إنقاذ الطفولة، 2003).

التحليل الذي قامت به إدارة تقييم العمليات بالبنك[11]، استخدمت مطابقة درجة الميل، والجمع بين بيانات من مناطق عمل المشروع مع بيانات من دراسة وطنية للتغذية أجرتها هيئة هيلين كيلر الدولية لتشكيل مجموعة مراقبة. هذا التحليل لم يجد أثر كبير للبرنامج على حالة التغذية ، رغم وجود أثر إيجابي على الأطفال الذين يعانون من سوء تغذية حاد .

قد تم التوقع أن يكون لمشروع التغذية المتكاملة ببنجلاديش أثر إيجابي على تأثير التغذية من خلال العديد من الافتراضات على طول السلسلة السببية، بعضها يظهر في الشكل (1).

[11] (OED ، الآن مجموعة التقييم المستقلة ، IEG)

المسألة الأولى هي ما إذا كان الناس في الواقع يعرفون عن البرنامج ويشاركون - العديد من المشاريع التنموية تسقط من الدور الاول حيث لا يتم بذل جهد كاف لشرح التدخل إلى المستفيدين المقصودين، أو إجراء تقييم واقعي للتكاليف والمنافع النسبية للمستفيدين. لكن مشروع التغذية المتكاملة ببنجلاديش قام بعمل جيد في هذا الصدد، وحوالي 90% من النساء المؤهلات جلبن أبنائهم، وإن كانت هناك بعض الاستثناءات الهامة، والتي يجب أن ينظر إليها فيما بعد.



الشكل 1 - السلسلة السببية لمشروع التغذية: نصائح غذائية والتغذية التكميلية

ثانياً، من الضروري أن يكون الأفراد المستهدفين هم المناسبون. استهدف البرنامج أمهات الأطفال الصغار. لكن فيما يتعلق بالصحة وتغذية الأطفال نجد أن الأمهات في كثير من الأحيان ليسوا صانعوا القرار، ونادرا ما يكونوا صانعوا القرار الوحيدين. بداية، المرأة في المناطق الريفية من بنجلاديش لا تذهب إلى السوق، بل الذي يقوم بالتسوق هو الرجل. أما في الأسر الكبيرة - حيث تعيش المرأة مع حماتها – كما يحدث مع عدد لا بأس من العائلات، تتولى الحماة مجال النشاط النسائي في الاسرة. في المناطق الأكثر محافظة من البلاد في الواقع، تكون معدلات المشاركة للنساء الذين يعيشون مع حمواتهم في المشروع هي أقل بكثير.

عند قدوم المرأة لوزن أطفالها، يتم ادخال الأطفال المستهدفين في البرنامج، وهم من يعانون من تعثر النمو أو سوء التغذية. لكن، أظهرت البيانات سوء استهداف كبير مع الأطفال من النوع الأول (لا يجري ادخال الأطفال

في البرنامج في الوقت المطلوب)، وأخطاء في النوع الثاني (ادخال أطفال لا يجب ادخالهم في البرنامج). لقد قمنا باختبار ممارسي التغذية المجتمعية ببعض الرسوم البيانية الخاصة بالنمو (تلك المستخدمة في التدريب)، واتضح أن معظمهم لم يستطع التعرف من خلال الرسوم البيانية بشكل صحيح أي الأطفال ينبغي قبولهم في البرنامج، وبالتالي حدث سوء استهداف. هذه مسئلة مهمة لأثر البرنامج، فقد وجدنا أن معظم الأطفال الذين يعانون من سوء التغذية استفادوا، وبالتالي كان يمكن أن يكون متوسط الأثر أكبر اذا كان البرنامج ركز على هؤلاء الأطفال، ولكن في الواقع كانت الموارد تذهب إلى أطفال لن يستفيدوا.

علاوة على ذلك، فبالنسبة للتغذية التكميلية ليكون لها أثرا مفيدا يجب أن تكون تكميلية ، في حين أنه في الواقع كان هناك تسرب على حد سواء (المواد الغذائية مقدمة إلى شخص آخر غير الشخص المقصود)، وكان هذا هو الحال بصفة خاصة بالنسبة للغذاء التكميلي المقدم للنساء الحوامل والاستبدال (تم أخذ المواد الغذائية بدلا من وجبة الطعام الطبيعية).

عودة إلى تغير السلوك، لقد حقق التواصل لتغيير السلوك تواصل، ولكن لم يحدث التغيير المطلوب في السلوك. المرأة التي تم إدخالها في البرنامج أصبح لديها معرفة أفضل بكثير عن "الممارسات الجيدة". ولكن كانت هناك فجوة كبيرة في تحويل المعرفة الى ممارسة: عدد كبير من النساء لم تحول هذه المعرفة إلى ممارسة. وكان السبب جزئيا القصور في الموارد: فالنساء في الأسر الفقيرة أقل احتمالا لتناول طعام أكثر خلال فترة الحمل، وتلك التي في الأسر التي لديها أرض أو تعيش مع قريب من الذكور كبير السن تكون أقل عرضة لأخذ مزيد من الراحة خلال فترة الحمل. ومرة أخرى إنها الحموات. وقالت إحدى المجوعات البؤرية لأحد الباحثين صراحة أنه "عند وفاة الحماة، ربما سوف نقوم بما تقولونه لنا، ولكن حتى ذلك الحين سوف نعمل بالطريقة التقليدية". أخيرا، من غير المرجح أن بعض تغيرات السلوك يكون لها أثر كبير على النتيجة النهائية لانخفاض الوزن عند الولادة، لا سيما تلك التي تهدف إلى زيادة الوزن اثناء الحمل (حيث أن وزن الأم قبل الحمل هو المؤثر).

باختصار، لقد ضعف أثر المشروع من خلال الحلقات الضعيفة والمفقودة في السلسلة السببية. بصفة عامة، لم يكن هناك أي أثر للمشروع. ولقد كان التحسن الذي أبدته بيانات رصد المشروع في الواقع يحدث في جميع أنحاء البلاد، وهذا هو السبب أن هيئة إنقاذ الطفولة لم تجد فرقا بين المشروع والمناطق المرجعية. في الواقع، إن التحسن كان راجعا إلى زيادة غلة الأرز، وارتفاع الدخل وانخفاض سعر الأرز، وليس مشروع التغذية المتكاملة ببنجلاديش.

ومع ذلك، فقد أشار التحليل إلى بعض الطرق الواضحة التي يمكن من خلالها تحسين أداء البرنامج: (1) مشاركة الحموات والأزواج في جلسات النصائح (المشورة) الغذائية، (2) تحديد وتشديد الاستهداف للبرنامج و(3) أداء أفضل للاستهداف عن طريق تدريب أفضل لممارسي التغذية في المجتمعات المحلية، وربما توظيف الممارسين بانتقائية. ومع ذلك، فإن العمليات الحسابية أظهرت ان هذا التدخل مكلف للغاية – فأي تدخل سيكون من الصعب أن ينفذ على نطاق أوسع في اطار محدودية الإدارة والموارد.

للأسف، لم تؤخذ دروس التقييم هذه في الاعتبار. كان فريق التغذية من البنك متشبث جدا بنموذج مشروع التغذية التكميلي ببنجلادش وتاميل نادو. وكان من المتصور ان ثبوت نجاحها في مشروع التغذية التكميلي بتاميل نادو (على الرغم من أنه لم يتم إجراء دراسة دقيقة بناء على معايير اليوم الصارمة)، وكان البنك أيضا يدعي نجاحا في بنجلاديش، وإن كان متنازعا عليه. المشاركة في هذه المناقشات أولا، ثم في وقت لاحق اصدار وثيقة من قبل فريق التغذية بالبنك أظهرت مشروع التغذية التكميلي ببنجلادش بمثابة نجاح بدون محاذير (البنك الدولي، 2006). وبناء على هذا الاعتقاد، اتخذ قرار طرح برنامج التغذية الوطني باستخدام نفس نموذج مشروع التغذية التكميلي ببنجلادش، رغم وجود أدلة تقييم لهذا النموذج الذي قد يكون قد نجح في تاميل نادو في حاجة الى بعض التكيف للعمل في بنغلاديش. وبعد ثلاث سنوات تهاوى برنامج التغذية الوطني وأغلق في وقت مبكر، وأرجأت دراسة الأثر المخطط لها لعدم وجود أثر واضح بسبب ضعف التنفيذ. وخلص

التقرير الختامي للبنك إلى توصيات لإصلاح البرامج خاصة المتشابهة بشكل ملحوظ لتلك التي تمت خلال العامين السابقين من خلال إدارة تقييم العمليات.

3. المبادئ

المبادئ الستة الرئيسية لتقييم الأثر القائم على النظرية هي ما يلي:

1 - وضع السلسلة السببية (نظرية البرنامج)

2 - فهم السياق

3 - توقع عدم التجانس

4 - تقييم دقيق للأثر استخدام الواقع المضاد الموثوق منه

5 - تحليل واقعي دقيق

6 - استخدم وسائل مختلطة

وضع السلسلة السببية (نظرية البرنامج)

تربط السلسلة السببية المدخلات بالنتائج والأثر. وهذا يعني أن السلسلة السببية تجسد نظرية البرنامج (أو نظرية التغيير) كيفية توقع حدوث الأثر المقصود نتيجة للتدخل. وهذه النظرية جزء لا يتجزأ من الإطار المنطقي التقليدي، على الرغم من أن هذا الاطار قد لا يركز على الافتراضات الأساسية، في حين أن اختبار الافتراضات أمر أساسي للنهج القائم على النظرية.

وثمة إنتقاد شائع لنهج السلسلة السببية وهي سلسلة خطية، وهذا يعني انها إما أحادية الاتجاه أو تتبع نهجا محدد مسبقا. انظر (وايت، 2009) لمناقشة المعاني المختلفة لكلمة "خطي" في سياق التقييم. ولكن الانتقادين ليسا صحيحين. في حين أنه قد يكون صحيحا أن مديري البرامج غالبا ما تصوروا إطارا بسيطا الى حد كبير يربط بين المدخلات والأنشطة والنتائج وتأثيرات وآثار وتقييم الاختبارات النظرية القائمة على الإفتراضات الكامنة

وراء هذه السلسلة من التفكير. واحدة من هذا القبيل من الافتراضات ان النتائج الملحوظة هي نتيجة لأنشطة المشروع ومخرجاته، وليس العكس. ثنائية الاتجاه السببية هو في صميم مناقشات تقييم الأثر: الانحياز في الاختيار الناجم عن قيام البرنامج بالاختيار الذاتي في البرنامج يعني على وجه التحديد متغيرات النتائج تؤثر على من يشارك ، وليس العكس. على سبيل المثال، المجتمعات التي ترتفع فيها مستويات رأس المال الاجتماعي هم أكثر عرضة لطلب الحصول على الأموال اللازمة لبرامج التنمية المجتمعية، التي تهدف إلى بناء رأس المال الاجتماعي، ولكن مجرد ملاحظة الفرق اللاحق في مستوى رأس المال الاجتماعي بين القرى المشاركة وقرى المقارنة، من المرجح أن يعكس الاختلافات السابقة على البرنامج أكثر من تأثير البرنامج .

وثمة إنتقاد أكثر قبولا هو أن هذا النهج قد يكون ثابتا إلى حد ما، في حين أن التدخلات يتم تبنيها وتتطور. تحمل النظم التي يتم وصفها في وثيقة المشروع القليل بالنسبة إلى الكيفية التي يتم بها تنفيذ البرنامج، إما لأنه قد تم إعادة تصميمها، أو لأن المديرين الميدانيين اعتمدوا على التفسيرات الليبرالية بدلا من إجراءات المشروع. في الحالة الأولى، ينبغي أن تعكس نظرية البرنامج التصميم الجديد، ووثيقة تقييم عملية التعلم التي أدت إلى هذا التصميم. في الحالة الأخيرة، نجد أن التناقض بين ما هو المقصود الذي ينبغي القيام به، وما يتم في الواقع هي مسألة التقييم الرئيسية: لماذا ظهرت هذه الإختلافات، وكيف أنها تؤثر على أداء البرنامج؟

مثال على تعلم المشروع يأتي من الدراسة التي أجريت على الصناديق الإجتماعية المذكورة أعلاه. انتقاد آخر للاستثمارات الممولة من الصندوق الاجتماعي هو أنها ليست مستدامة اذ لا يرصد مبالغ لعمليات التشغيل والصيانة (O&M). تستخدم الصناديق الاجتماعية في الأصل لجنة مركزية للموافقة على جميع الطلبات، بفرض أن الوزارات الممثلة في هذه اللجنة قد وافقت وتعهدت على تغطية النفقات التشغيلية عندما وافقت على المشروع. ولكن هذا النظام لم يعمل، وبدأت الصناديق الاجتماعية في الدخول في "الاتفاقات الشاملة"، لتغطية جميع المشاريع لكل وزارة. هذا النظام أيضا لم ينجح بالكامل، مما جعل بعض الصناديق الإجتماعية تسعى للحصول على موافقة الوزارة المعنية على أساس كل حالة على حدة، والبعض الآخر يطلب الخطط المحلية

المستدامة، وبعضها الآخر يرصد جانبا من الموارد في صندوق للصيانة (انظر البنك الدولي، 2002).

يجب أن تكون نظرية البرنامج ديناميكية وتتيح التعلم من الميدان، والذي هو إعادة تأكيد للحاجة إلى الربط ما بين النظرية والبيانات. تتعامل نماذج النهج المستند إلى التحليل الإحصائي على النموذج كما هو وارد، وتختبر ببساطة مدى تناسب جودة البيانات مع النموذج. والممارسون لهم طرق مختلفة لضمان صلاحية البيانات للغرض، كما هو الحال في بيان كوس بأن البيانات سوف تعكس ما اذا كان قد تمالتعامل معهم لفترة طويلة بما فيه الكفاية (اقتبس من قبل ليمر، 1983). ومع ذلك، فإن تحليل البيانات يتيح لها قيادة النظرية، والبحث عن اتجاهات في البيانات. هذا النهج يبدو غير منظم، ولكن بالطبع لا يمكن أن تكون الممارسة الإحصائية خالية من النظرية، حيث أن النظرية هي التي تحدد في المقام الأول ما هي البيانات التي يتم جمعها وتحليلها، بل ينبغي أن تكون مستعدة لنظرية التكيف مع المفاجآت في البيانات . وهذا النهج قد يبدو أقرب إلى استخراج البيانات، ولكن الواقع مختلف تماما. الباحث عن البيانات يعرف ما يبحث عنه ويحفر حتى العثور عليها. ومحلل البيانات، من ناحية أخرى، يسعى من خلال البيانات إلى إظهار الاتجاهات، المتوقعة أو غير المتوقعة، (انظر موخرجي وآخرون، 1998، الفصل 1 للاطلاع على مناقشة أوفى).

ربما آخر انتقاد صحيح لهذا النهج هو أنه من خلال التركيز على السلسلة السببية للدراسة سيغيب ملاحظة الآثار غير المقصودة. ويمكن تجنب هذا الضعف بطريقتين. الأولى، يمكن لتطبيق دقيق لنظرية برنامج تحديد العواقب غير المقصودة المحتملة، على سبيل المثال من خلال التفكير في الآثار البيئية، والتي قد تم التعامل معها بطريقة سطحية بواسطة مصممي البرامج. الثانية، العمل الميداني الأولي، بما في ذلك التحليل بالمشاركة، وهو جزء مهم من تصميم التقييم، والذي يمكن أن يلتقط نتائج غير مقصودة من هذا القبيل، ومن ثم يمكن إدراجها في إطار التقييم.

ويرتبط أيضا بمسألة الآثار غير المقصودة "نظرية من؟" ان نظرية التصميم الجيد المبني على النظرية سوف

يأخذ بعين الاعتبار النظريات المتنافسة في كيفية وضع برنامج عمل. ومديري البرامج لديهم رؤية واحدة، ولكن الموظفين الميدانيين، والمستفيدين أو المعلقين الآخرين قد يكون لهم وجهات نظر مختلفة تماما. على سبيل المثال، مشاريع الصندوق الاجتماعي (صرف الأموال المخصصة للتنمية على مستوى المجتمع المحلي) حيث أكد مديري البرامج أن لها آثار ايجابية على التنمية المؤسسية على المستوى المحلي والوطني من التعلم عن طريق مثال (رؤية ما فعله الصندوق الاجتماعي)، والتعلم عن طريق العمل حيث أن هذه الوكالات تشارك في التنفيذ مع الصندوق الاجتماعي. ومع ذلك، يذكر النقاد أن الصناديق الاجتماعية تجاوز الإجراءات الحكومية القائمة، مما يقوض بشكل مباشر هذه المؤسسات القائمة (عن طريق انتداب الموظفين)، وأقل مباشرة من خلال تعطيل التخصيص الأمثل للموارد من قبل الوزارات المعنية. وبالتالي يراعي التقييم كل من نظرية البرنامج الرسمي، والتنافس المضاد من قبل الوزارات (انظر البنك الدولي، عام 2002، للدراسة الكاملة - كارفاليو وآخرون، 2002، للاطلاع على موجز - وكارفاليو ووايت، 2004، لعرض النهج القائم على استخدام النظرية المستخدم).

نقطة البداية المعتادة لوضع نظرية البرنامج هي وثائق المشروع. إذا كان هناك إطار منطقي (الإطار المنطقي) فان هذا الإطار يجسد نظرية البرنامج. ومع ذلك، فمن غير العادي اعداد وثيقة مشروع تشمل جميع الافتراضات الأساسية، على الرغم من أن بعضها قد تظهر "كمخاطر". والخطوة التالية هي ادارة نظرية البرنامج المقترح من قبل مديري البرامج، حتى لو لم يكن قد تم التفكير فيها بشكل واضح من قبل، سيكون لديهم وجهات نظر حول أي وثيقة من هذا القبيل التي يتم انتاجها. هذه العملية هي فرصة جيدة لإشراك مديري البرامج، مما يتيح لهم التأثير في تصميم التقييم بطرق مفيدة[12]. وتتمثل الخطوة الثانية لقراءة دراسات التقييم الحالية والكتابات الأكاديمية، إن وجدت، في التدخل الذي يجري تقييمه أو البرامج المماثلة، والتي سوف تتيح التعرف على الروابط الضعيفة في السلسلة السببية. على سبيل المثال، الخطأ في الاستهداف مشكلة كثيرا ما يستشهد بها ، وخاصة في برامج تمويل المشاريع الصغيرة (على سبيل المثال موسلي وهولم، 1996). وهناك

[12] الرد المعتاد لمديري البرامج هو أنه ليس الوقت المناسب لتقييم البرنامج لأنه قد تم إعادة تصميمه ، أنه لديهم الدراسات الخاصة بهم ، هناك تغيرات في الحكومة ، الوزير ، أو مدير المشروع ... الخ. لابد التغاضي عن مثل هذه الاعذار بصورة مناسبة وعدم محاولة التأثير على المخرجات. ولكنه له معنى ان يتم الاستفادة من ايمان مديري البرامج من حيث اعتقادهم لما هية الاسئلة المهمة للتقييم.

نقطة أكثر دقة وهي أنه في تمويل المشاريع الصغيرة للسيدات، اذ يمكن أن تستخدم القروض من قبل أفراد الأسرة من الذكور، الأمر الذي يؤثر على التأثير النهائي المتوقع مثل صحة الأطفال وتغذيتهم . وجهة النظر هنا ان يدمج أولئك العاملون في الميدان والمستفيدون. من المفيد لأي مقيّم أن يسأل نفسه: كيف سيحكم القروي العادي على هذا المشروع؟ وكيف يصل إلى معرفة ما إذا كان ذلك ناجحا؟ ولماذا يشاركون؟. ومن المفيد أن تتم المحاولة، على الرغم من أن علم أنثروبولوجي التنمية قد علمنا أن وجهات النظر المحلية حول المشاريع قد تكون مختلفة جدا عن ما هو متوقع، بسبب اختلاف التصورات والإحتياجات أو فشل التواصل من قبل العاملين بالمشروع.

فهم السياق

فهم السياق أمر حاسم لفهم تأثير البرنامج، وبالتالي تصميم التقييم. يعني السياق الظروف الاجتماعية، والوضع السياسي والاقتصادي الذي يتم في اطاره تنفيذ البرنامج، والتي يمكن أن يؤثر على الطريقة التي تتم بها السلسلة السببية. ان مردود البرامج المتشابهة يمكن أن يختلف في سياقات مختلفة: كما هو الحال مع نموذج ناجح TINP فهو على ما يبدو لا يعمل خارج السياق كما في بنجلاديش. ومع ذلك، فإن البرامج المتشابهة، هي شيء من المثالية، اذ نادرا ما تتحقق نفس التأثيرات في الظروف الميدانية المتباينة، والتي هي في حد ذاتها تشكل جزءا هاما من السياق. وعلاوة على ذلك، فان فهم السياق يساعد على توقع عدم التجانس، وسوف يساعد أيضا عند التعميم كما سيأتي فيما بعد.

يعني فهم السياق قراءة متأنية لوثائق المشروع قبل الشروع في تصميم التقييم، ولكن يعني أيضا التعرض لقراءات أكثر في الانثروبولوجيا والإقتصاد السياسي، كما سيناقش في إطار استخدام الوسائل المختلطة.

يساعد فهم السياق ايضا في التعميم. وأظهرت دراسات عن مشروعات البنك الدولي لدعم التعليم الأساسي في غانا، وصحة الأم والطفل في بنجلاديش قصص النجاح الشامل. في غانا، ساهم إعادة تأهيل المدارس وتوفير

الكتاب المدرسي بدرجة كبيرة في تحسين معدلات الالتحاق ومخرجات التعلم (البنك الدولي، 2004). كان هناك اثنان من الجوانب السياقية وراء هذه النتيجة. الأول، هو أنه، وبعد سنوات من الأزمة في النظام المدرسي - وكان في حالة مزرية جدا في الواقع، مع عدم كفاية البنية التحتية، واللوازم المدرسية تكاد تنعدم – فقد كان لتجديد المدارس وتوريد الكتب المدرسية تأثير في هذا السياق الذي كان لا يمكن أن يحدث لو لم يتم هذا التدخل وتعمل المدارس بشكل جيد نسبيا. الثاني، كان هناك دعما سياسيا قويا للبرنامج، مما ساعد على سهولة التنفيذ (البرنامج كان جزءا من إصلاح أوسع للعملية التعليمية). وكان التزام الحكومة أيضا عنصرا أساسيا في نجاح التخطيط الممول من المساعدات، التي أسفرت عن عملية الانتقال الديموغرافي المتسارع في بنجلاديش، مع انخفاضات كبيرة في معدلات الوفيات والخصوبة (البنك الدولي، 2005). تحولت البلاد من دولة بدون مرافق بعد الاستقلال مباشرة ، ويجري وصفها كحالة متتالية من المجاعات المتلاحقة ، إلى وجود دولة ذات نظام لامركزي للصحة وتنظيم الأسرة، وصولا الى تسليم خدمات وسائل منع الحمل، لمدة عشر سنوات. قد تتعثر البرامج المتشابهة الطموحة، إذا كانت الحكومة ليس لديهم الرغبة في أن ترى هذه النجاحات.

توقع عدم التجانس

يساعد فهم السياق على توقع امكانية عدم تجانس الأثر. يمكن للمردود (وهذا هو تأثير التدخل) أن يختلف وفقا لتصميم التدخل، أو سمات المستفيدين أو الظروف الإجتماعية والإقتصادية. ويمكن من خلال إختبار النظرية الكامنة وراء عدم التجانس. توقع عدم التجانس المحتمل هام لسببين: أولا، يحتاج إلى حسابات لحجم العينة لتعكس مستويات الإختلاف التي سيتم استخدامها في التحليل: كلما زادت درجة الاختلاف تكون العينة المطلوبة أكبر (لكل من العينة المعاملة وغير المعاملة). ثانيا، الاحتمال البسيط يشير إلى أنه إذا كنا سنقوم باختبار الأثر في 20 مجموعة فرعية، وبشكل عشوائي، فإننا سنجد لها تأثير كبير في واحدة من تلك على مستوى 5%. تتطلب الممارسات الجيدة، اللازمة للمعاملات الطبية العشوائية أن يتم تعريف المجموعات الفرعية لفحصها من قبل جامعي البيانات. يساعد النهج القائم على النظرية في التحديد المسبق لمثل هذه الجماعات، ويقدم تفسيرا مقبولا للفروق في الأثر. ومع ذلك، فإنه يجب الأخذ فيالاعتبار طول العملية ما بين النموذج النظري والجمع

العلمي للبيانات.

عند النظر في برامج تغذية الأطفال، نجد أن الأطفال الذين يعانون من سوء التغذية هم أكثر عرضة للتفاعل من خلال زيادة الوزن أكثر من الأطفال الذين يحصلون بالفعل على تغذية جيدة، الا أن الأطفال الذين يحصلون على تغذية سيئة، يمكن أن يصابوا بالاسهال الذي يمنع التغذية الفعالة وزيادة الوزن. والبرامج ذات الاستهداف الأفضل، بالتالي، يكون لها أثر متوسط، وهذا الأثر سيظهر أكبر في موسم الجفاف - كما وجد بالفعل أن هذا هو الحال مع مشروع التغذية المتكاملة ببنجلاديش. الأطفال الأصغر سنا، من المرجح أن يستفيدوا أكثر، والأطفال الذين يعانون من توقف النمو في مرحلة الطفولة لن يتأثر طولهم بدرجة ملحوظة نتيجة تحسين التغذية في السنوات اللاحقة . تظهر المكاسب الراجعة الى تحسين التغذية لدى الأطفال الأصغر من ثلاث سنوات من العمر. وبالتالي، يختلف الأثر حسب عمر الفئة المستفيدة، والحالة التغذوية قبل التدخل، والتي تتأثر موسميا. يمكن أن يختلف الأثر بحسب الوضع الاجتماعي والاقتصادي، على سبيل المثال، الاستبدال (إستخدام "التغذية التكميلية" لتحل محل الوجبة الغذائية الحالية) والذي يرجح حدوثه اكثر في الأسر الفقيرة.

لهذه الأسباب قد تحول الاتجاه في برامج التغذية بعيدا عن التغذية المدرسية الى استهداف من هم دون سن الثالثة، كما هو الحال في البرنامج ببنجلاديش الذي نوقش بالتفصيل أعلاه. ولكن لا يزال يمكن أن نتوقع للتغذية المدرسية أن تسفر عن تحقيق مكاسب في التعلم. إن نقص السعرات الحرارية يجعل الأطفال يصابون بالتعب والفتور، لذلك فإن برنامج للتغذية يجعلهم أكثر انتباها في الفصول، مع التحذير من أن معظم الناس يصابون بالنعاس بعد تناول وجبة جيدة، لذلك فان التوقيت هام هنا. ولكن لتوضيح الأمور، فإنه لتحقيق مكاسب التعلم للأطفال يجب أن يكونوا أكثر إنتباها . وهناك إفتراض حاسم بالنسبة لجميع التدخلات وهو أن يتم معالجة المشكلة الحقيقية. ولا فائدة للطفل اليقظ إذا كان المعلم غائب، وعلى الأرجح سوف يتعلم أقل إذا لم تكن هناك مواد تعليمية. لذلك، يمكن توقع أثر اكبر لبرامج التغذية في المدارس ذات الأداء الجيد من تلك التي تفتقر للتجهيزات، والتي يغيب فيها المدرسين. أشياء مماثلة فيما يتعلق بالتحويلات النقدية المشروطة، التي تزيد من

الطلب على التعليم، ولكنها قد لا تحسن نتائج التعلم، أو حتى الالتحاق، إذا كانت هناك معوقات من جانب العرض (رافاليون، 2009).

وثمة جانب آخر من عدم التجانس والتكامل بين التدخلات الممكنة، على سبيل المثال تمويل المشاريع الصغيرة له تأثير كبير إذا كان مصحوبا بخدمات دعم الأعمال التجارية. أو ربما أثر عكسي، حيث أن أثر الاثنين مجتمعين هو أقل من مجموع الإثنين كل على حدة . التصاميم التي تستكشف أوجه التكامل ذات أهمية لوضع السياسات.

يمكن أن يختلف الأثر أيضا عبر الزمن، على الرغم من إفتراض (ضمني في كثير من الأحيان) أن للأثر مسار خطي (وولكوك، 2009). يختلف الأثر الخطي عن الإنتقادات التي سبق مناقشتها حول السببية أو أحادية الإتجاه لكونه نهج ثابت لتنفيذ برنامج في بيئة متغيرة بإستمرار. حتى عندما يكون تصميم البرنامج لم يتغير، وقد تم تأسيس اتجاه السببية، فان أثر التدخل يتغير مع مرور الوقت، وستكون النتائج حساسة جدا لهذه النقطة في الوقت الذي يتم فيه قياس الأثر. على سبيل المثال، بالنسبة للمشاريع التي تسعى لزيادة المشاركة والتمكين للفئات المهمشة، تشير الأدبيات إلى أن الشكل الأكثر إحتمالا لتأثير مثل هذه المشاريع مع مرور الوقت هو منحنى J، أي أن الأمور تزداد سوءا قبل أن تتحسن. هذا هو أحد المجالات التي لم يتم إستكشافها بشكل كاف بإستخدام تقييم الأثر المرتكز على النظرية، ولكنها حالة يجب أن يُفسح المجال لدراستها بشكل خاص. في حالة مشروع التغذية المتكاملة ببنجلاديش الذي نوقش في وقت سابق، قد يتسبب في إحداث صراعات أولية بين النساء وأزواجهن، والحموات بسبب الوعي المتزايد من النساء، الأمر الذي يفسر عدم ظهور أي آثار للتغذية، ولكن ربما بمرور الوقت يكشف التقييم عن تأثير إيجابي نظرا للتغييرات الإجتماعية الأوسع نطاقا، وزيادة تحسن وضع المرأة في المناطق الريفية من بنجلاديش.

ويرتبط تحديد عدم التجانس بامكانية التعميم. أظهرت تجربة عشوائية محكمة في كينيا وجنوب أفريقيا وأوغندا

عن تأثير ختان الذكور على انتقال فيروس نقص المناعة البشرية / الإيدز، ان الرجال المختونين كانوا أقل إحتمالا للإصابة بالمرض بدرجة كبيرة (انظر، على سبيل المثال، واور وآخرون، 2008، عن أوغندا). وكان أحد جوانب عدم التجانس هو السن. وينبغي أن يكون هناك امتناع عن ممارسة الجنس لمدة شهر بعد الختان حتى يمكن التئام الجرح؛ اذ أن ممارسة الجنس في تلك الفترة يعرض للاصابة، بدرجة أكبر. تنفيذ الإجراء على الفتيان فى عمر 12 عاما لا يحمل هذه المخاطرة وهي التعرض للاصابة خلال شهر بعد الختان. لكن الاكبر سنا من الذكور غالبا ما يكونون غير قادرين على الإمتناع عن الجنس لمدة شهر كامل ، وبالتالي يقل التأثير الايجابي للختان. ومع ذلك، وجدت الدراسات إنخفاض في مخاطر إنتقال المرض ما بين 30% وأكثر من 50% نتيجة للختان. ويمكن تعميم هذا المستوى من التأثير فقط للسكان ذوي الأنماط المماثلة من السلوك الجنسي. في المجتمع الذي يمارس فيه الرجال الإمتناع عن ممارسة الجنس، أو الممارسة مع شريك واحد أو إستخدام الواقي الذكري، فانه لن يكون هناك أي تأثير ناتج عن التدخل .

الأثر (المردود)

يستخدم التقييم الدقيق للأثر ليقيس أيضا التغيير الذي يحدث في المجتمع بدون التدخل وهو بالطبع عنصرا رئيسيا من عناصر تقييم الأثر المرتكز على النظرية. ويتم ذلك من خلال مجموعة لا تتعرض للتدخلات، والتي يجب ان تحدد بطريقة تتجنب التحيز في الإختيار، بإستخدام النهج التجريبي أو شبه التجريبي. يساعد وجود البيانات على تعزيز التصميم، لذلك فان تجميع بيانات خط الأساس يجب أن يصمم بطريقة تسمح بإعادة تحديد أسر العينة – وحين لا تتوفر هذه البيانات، لابد من توفيرها بإستخدام مجموعات البيانات المتاحة أو الاستدعاء، على الرغم من ضرورة الحاجة لتوخي الحذر مع الحالة الاخيرة (انظر بامبيرجر، 2009). بالإضافة إلى الانحياز في الإختيار، فهناك قضايا مهمة في التصميم وهي إمكانية تسرب بعض الآثار (من خلال تأثير التدخل على مجموعة المقارنة)، أو تدخلات أخرى (التأثير على مجموعة المقارنة من قبل التدخلات الأخرى) .

التحليل الواقعي الدقيق

يحتاج تحليل الأثر من خلال قياس التغيير الذي يحدث بدون تدخل إلى أن يستكمل من خلال تحليل واقعي صارم من مختلف الأنواع. يعتمد العديد من العلاقات في السلسلة السببية على تحليل واقعي. في حالة مشروع التغذية المتكاملة ببنجلاديش. يشمل هذا ضعف الإستهداف وأسباب ذلك، وتحديد حالات التسرب، وحقيقة أن تحسين المعرفة لم يتحول إلى ممارسة عملية.

يجب أن يتم تحليل الإستهداف من خلال تحليل واقعي لبيانات العينة الممثلة، أن تكون جزءا من معظم، إن لم تكن من الكل، ودراسات الأثر: من المستفيد من هذا البرنامج؟ إلى أي حد هناك مجموعة مستهدفة محددة، ثم ما هو مدى الأخطاء التي تحدث، ويمكن تحديدها كميا وتحديد مصدرها، كما حدث في بنجلاديش. وينبغي أن يتم تحليل الاستهداف على مستويات مختلفة. وجد في حالة الصناديق الإجتماعية أن إستخدام خرائط الفقر يعني أن تركز الصناديق الإجتماعية في العديد من البلدان على الأحياء الأكثر فقرا، ولكن داخل هذه المناطق فقد كانت المجتمعات الأكثر غنى أكثر قدرة للوصول إلى موارد المشروع (البنك الدولي، 2002). في حالة كهربة الريف، كانت المجتمعات الأفضل حالا هي الأكثر قدرة على توصيل الكهرباء، ولكن الأسر الفقيرة في المجتمعات المحلية تبقى بدون وصلات كهربائية لسنوات عديدة بسبب عدم قدرتهم على تحمل رسوم ادخال الكهرباء (البنك الدولي، 2008).

يتم تحليل الإستهداف من خلال بيانات ممثلة. من خلال جمع بيانات تسمح بالتحيز في الإختيار سيكون تحليل الأثر ليس ممثلا للمجتمع ككل، وبالتالي لا يمكن إستخدام هذه البيانات للإجابة على سؤال مثل "ما هي النسبة من 20% الافقر من السكان حققت إستفادة من المشروع؟" الا اذا كان هناك أوزان لأخذ العينات ليمكن أخذ عينة ممثلة.

ونقطة ثانية فيما يتعلق بالإستهداف هي أنه ممارسة ثنائية التغير، الأمر الذي يتطلب إيجاد علاقة بين المشاركة والخصائص ذات الاهتمام (هذه الخصائص قد تكون خصائص الفرد أو الأسرة أو المجتمع). ويتطلب النهج

شبه التجريبي إستخدام التحليل متعدد المتغيرات لتحديد المشاركة، لكنه عموما من الخطأ إستخدام هذه النتائج لتحليل الإستهداف في المشروع، الذي ينبغي أن يعتمد على إحصاءات وصفية. ما إذا كان البرنامج يصل إلى أسفل 20% هو بيان بناء على جدولة ثنائي المتغير، وليس دلالة إحصائية لخمس السكان في الإنحدار متعدد المتغيرات. ما الذي يمكن أن يقوم به التحليل هو تسليط الضوء على العوامل التي تفعل المشاركة، وبالتالي تساعد على تفسير نتائج تحليل الإستهداف ثنائي المتغير. على سبيل المثال، قد يكون التحليل متعدد المتغيرات لمشروع في الهند يظهر انخفاض مشاركة سكان القبائل، التي هي من بين الأكثر فقرا في بعض مجالات البرنامج، وهو ما يفسر ضعف عملية الإستهداف.[13]

مثال لحالة إستخدام غير كامل للتحليل الواقعي هو إختبار ما إذا كان الأفراد الذين تعرضوا للتدريب تعلموا ووضعوا ما تعلموه موضع التطبيق، وهو المنشود من التدريب. وأظهرت الدراسة أن الأمهات في مشروع التغذية المتكاملة ببنجلاديش إكتسبن المعرفة، ولكن الكثيرات لم يضعنها موضع التنفيذ. وكذلك يمكن لممارسي التغذية المجتمعية إجراء عمليات الوزن، ولكن بشكل حاسم، لم يتعلموا تفسير الرسوم البيانية للنمو بشكل صحيح. لم يتم القيام بمثل هذا التحليل في كثير من الأحيان، ولكن هناك مجالا كبيرا لذلك. هل المعلمون المدربون يعرفون طرق التدريس المحسنة ويضعوها موضع التطبيق؟ وترى دراسة للبنك الدولي في غانا أن العديد منهم لا يفعل ذلك.[14]

كما هو الحال في حالة بنجلاديش، يمكن للتحليل الواقعي، في كثير من الأحيان، تسليط الضوء على فاصل حاسم في السلسلة السببية، وبذلك يوضح لماذا الأثر المنخفض للتدخل. وجدت دراسة أخرى لـOED أن التدريب وزيارات الإرشاد الزراعي للمزارعين في كينيا لها تأثير على تحقيق المردود. في الأساس، المشروع يمول البحوث الزراعية الجديدة في مراكز البحوث، والنتائج المفيدة يتم تمريرها الى العاملين في مجال الإرشاد

[13] مصطلح الثروة مهم أن يكون أكثر تحديدا عندما يتم استبعاد دمية القبلية، ولكنه يصبح غير مهم عندما يتم تضمين المتغير الأخير. وبالتالي فإنه هو الوضع القبلي في حد ذاته الذي يقود المشاركة وليس الفقر. هذا النهج قد لا يكون من الممكن تطبيقه دائما على حساب الدرجة العالية من التعاون الخطي بين المتغيرات التفسيرية المحتملة، مثل تلك المذكورة والتعليم والمكان.

[14] أن الملاحظة الفصول الدراسية تكون أفضل طريقة لقياس الممارسة، ولكن تم استبعادها بسبب التكلفة. قد يعتقد المرء أن مجرد سؤال المعلمين سيكون متحيزا حيث أنها ستقدم تقريرا باستخدام أساليب محسنة حتى لو لم يفعلوا ذلك، ولكن ممارسات أثبتت أن المعلمين إما اظهروا الاستغراب، أو أن عدم معرفتهم ادى الى أنهم لم يعرفوا اي هي "الإجابات الصحيحة".

ومن ثم للمزارعين. نجد في الواقع، أن الدروس المستفادة من هذه البحوث لم تعرض على العاملين في مجال الإرشاد، الذين كانوا يقدمون الرسائل الارشادية للمزارعين لتبني ممارسات معظمها تم تبنيها بالفعل منذ فترة طويلة (البنك الدولي، 2000).

ومع ذلك، في بعض الأحيان فإن ما يبدو شرطا لإجراء تحليل واقعي يمكن في الواقع أن يتطلب تحليل الواقع في غياب التدخل. في حالة توفر منح دراسية يكون الهدف من المنحة المدرسية زيادة عدد المسجلين، وتحسين نتائج التعلم على حد سواء. ولكن كيف يتم ذلك؟. يجب أن يعطي التفسير توضيح حول إستخدامات الاموال. قد يبدو هذا تحليلا واقعيا واضحا - تتبع إستخدام الأموال: التحقق مما وصل بالفعل إلى المدارس، وكيف يتم الإنفاق عليها، وكلاهما في الواقع أجزاء مفيدة لهذه الدراسة. لكن، إذا كانت المدارس لديها بالفعل بعض الموارد تحت تصرفها فهناك إمكانية للتبادل. ويمكن لتحليل من قبل مقابل تحليل من بعد لأنماط الإنفاق أن يسفر عن الواقع في غياب التدخل في هذه الحالة، على الرغم من أن مقارنة المدارس التي حصلت على منح وبالتي لا تحصل عليها من حيث تحسين حالة المدرسة وتوفير المواد اللازمة من المرجح أن يكون التصميم أقوى.

الطرق المختلطة

"الطرق المختلطة"هي استخدام مزيج من الأساليب النوعية والكمية في تقييم واحد. جميع الدراسات الكمية لديها قدر من التحليل النوعي - على الأقل قراءة وثائق المشروع – لذلك، هي مسألة تتعلق بمدى إستخدام كل من التحليلات الكمية والنوعية.

تأتي الدعوة إلى إستخدام طرق مختلطة عموما من مؤيدي النهج النوعي. ولكن، في دراسات التنمية، نجد أن التحليلات النوعية هيمنت على التقييم حتى وقت قريب جدا، لذلك نجد أن خطوة رئيسية نحو إستخدام الأساليب المختلطة هو في الواقع زيادة استخدام الأساليب الكمية في الدراسات النوعية. إلا انني اود أن اشير الى زيادة

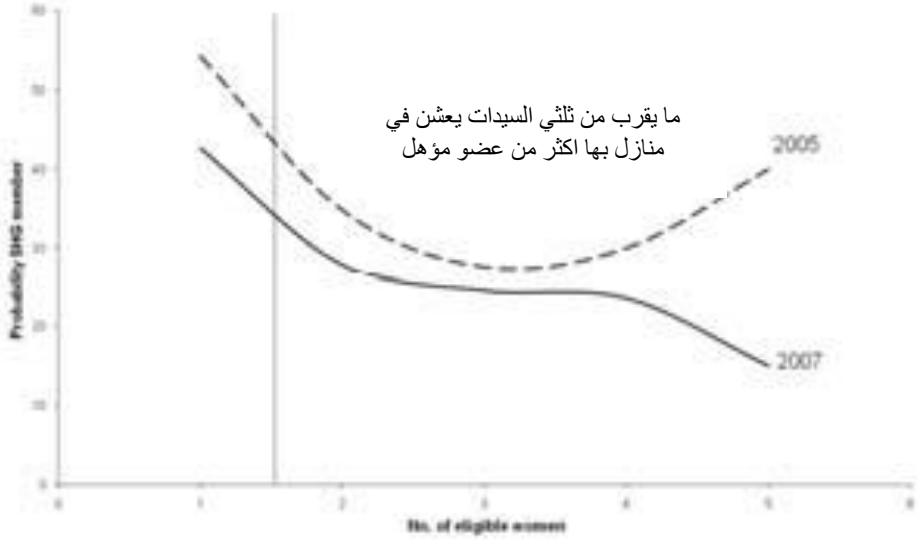
استخدام البيانات النوعية في الدراسات الكمية، وهي قضية تعاملت معها بإستفاضة في وايت (2008). وسأقدم هنا ثلاث نقاط عامة.

أولا: إستخدام البيانات النوعية يعني مجموعة واسعة من الأنشطة، وليس مجرد الإعداد لبعض مجموعات التركيز (من وجهة نظري من أضعف أشكال جمع البيانات النوعية، إلا إذا تمت بدقة). وهي تشمل، على سبيل المثال، قراءة الأدبيات الأنثروبولوجية والسياسية في سياق التدخل لإعداد تصميم التقييم. في حالة بنجلاديش، جاء تحديد تأثير الحموات، من خلال قراءة الأدبيات الأنثروبولوجية (خصوصا وايت، 1992). أدت هذه الرؤية الى تعديل المقطع الخاص بخصائص الأسرة من الإستبيان للتعرف على تلك المرأة التي تعيش مع حماتها (على سبيل المثال: زوجة ابن رب الأسرة وتعيش مع حماتها، زوجة رب الأسرة؛ وتعيش فيها والدة رب الأسرة، حماتها؛ واخت زوجة رب الأسرة وتعيش معهم أم رب الأسرة أيضا)، وبذلك يمكن القيام بالتحليل الكمي في ضوء المعلومات النوعية.

تبدأ التقنيات المسخدمة في تقييم الأثر من "سياحة التنمية" (قضاء يوم أو نحو ذلك في الميدان) ، مرورا بالتحليل السريع بالمشاركة، إلى تضمين عالم الأنثروبولوجيا في منطقة المشروع، وهذا الأخير هو نهج غير مستغل جيدا، ويمكن إستخدامه في الدراسات طويلة الأجل. ثانيا هي أنه، على الرغم من السخرية من سياحة التنمية إلا أنها تشكل جزءا أساسيا من تقييم الأثر المرتكز على النظرية. فليس هناك بديل لقضاء بعض الوقت في الميدان، وأنه من الصعب أن تعرف كيف يمكن تحليل البيانات بشكل معقول من دون التعرف على موقع العمل. يؤدي قضاء بضعة أيام فقط في الميدان الى التعرف على إسلوب تنفيذ المشروع ومجموعة من الأنشطة - وليس فقط تلك التي إختارها العاملون بالمشروع وسوف يساعد على حد سواء في تصميم وتنفيذ الدراسة. ومن المفيد أيضا القيام بزيارة المنطقة خارج مناطق المشروع .

يمكن أن أعطي أمثلة كثيرة لرؤى من مناقشات مع موظفي المشروع والمستفيدين وأصحاب المصلحة

الآخرين. سأعطي إثنين فقط: الأول من تقييم لمشروع سبل المعيشة الريفية التي شملت إعطاء القروض من خلال مجموعات المساعدة الذاتية النسائية. رجل واحد اشتكى من أن إبنته غير المتزوجة (22 سنة) لم تتمكن من الحصول على قرض نظرا لأن زوجته تلقت بالفعل قرضا. هذا التعليق أدى إلى معرفة أن القرويين إعتبروا القرض للأسرة لا للفرد، الأمر الذي دفع معدلات مشاركة أقل في الأسر التي لديها أكثر من امرأة واحدة مؤهلة للانضمام الى مجموعة المساعدة الذاتية (شكل 2). الهدف من المشروع هو أن جميع النساء المؤهلات ينبغي أن تشارك، ولكن هذا ليس هدفا واقعيا طالما العائد يعود على الاسرة، وليس الفرد. المثال الثاني يبين كيف أن حسن الإختيار يمكن أن يكون نقطة فاعلة. في العمل الميداني لتقييم الصندوق الاجتماعي في زامبيا كان لافتا كيف أن الجميع - من المديرين وإلى موظفي البرنامج، إلى القرويين - يقولون "المجتمع" إختار هذا المشروع، على الرغم من أنه كان واضحا أنه في الواقع تتم عملية أكثر انتقائية (انظر وايت وفاجا، 2008، لمناقشة أطول). وفي الحقيقة أن "المجتمع" هو في الواقع التصور الضيق، وهو يعني لجنة المشروع ، والتي تدار من قبل ضابط البرنامج الإقليمي الذي قال لنا عند الرد على هاتفه المحمول: "يجب أن أذهب الآن، لدي المجتمع في مكتبي".



شكل 2: معدلات المشاركة فى البرنامج وفى مجموعات المساعدة الذاتية في ولاية اندرا براديش مقارنة
عدد السيدات المؤهلات
بعدد النساء المؤهلات في المنزل

المصدر: بيانات المسح IEG

وحيث أن البيانات المتاحة تسمح بتقييم الأثر بدون الحاجة الى جمع بيانات جديدة (ذلك أمر ينبغي تشجيعه، حيث أننا على استعداد دائما لجمع بيانات جديدة، بينما البيانات الموجودة غير مستغلة بالدرجة الكافية)، فإن هناك خطرا من الباحثين بإجراء دراسات لتحديد الأثر دون التعرض للتدخل. مثل هذه الدراسات تفتقر الى الإرتباط مع السياسات، نظرا لسوء فهم كيفية إحداث المردود بواسطة التدخل.

أخيرا، يجب أن تسمح الميزانية ببعض الأنشطة ذات الطبيعة البحثية، حيث يتم متابعة الألغاز في البيانات مع إجراء اعمال ميدانية إضافية. إستخدام مجموعات التركيز لمعرفة الأسباب التي أدت إلى فجوة الممارسة في بنجلاديش هي مثال على مثل هذه الأعمال. مثال آخر يأتي من الدراسة سابقة الذكر عن تمويل مجموعات المساعدة الذاتية في ولاية اندرا براديش في الهند. كان لدينا بيانات، وشملت الدراسة معيار قياس مستويات المعيشة وتأثيرها على الوحدات الإقتصادية المنزلية. وأظهر تحليل هذه البيانات إنخفاض العائد في معظم الوحدات ، بما فيها أقلية مؤكدة ذات الأنشطة الخاسرة. إلا أن البيانات الموجودة في الواقع لم توفر وسيلة لفهم كيف تستطيع هذه الوحدات أن تظل عاملة. وبالتالي، فقد قمنا بتكليف ما يمكن أن أسميه""الاثنوغرافيا الكمية" لإعادة زيارة جميع الأسر التي تم مسحها، وإستخدام استبيان شبه منظم لتحديد التدفق النقدي اليومي للوحدات، ومدخلات العمل من أفراد الأسرة (والموظفين، رغم أن هذه كانت نادرة جدا). وأكدت النتائج، في الواقع، إنخفاض مستوى الدخل من هذه الأنشطة (20-30 روبيه في اليوم الواحد لم يكن غير عادي، مقارنة بالأجر اليومي للشخص 50-70 روبيه)، بالاضافة الى طبيعة المخاطرة في العديد من هذه الأنشطة (موت الماشية، وخاصة من الماعز، وعدم كفاية حجم السوق).

4. تقييم الأثر القائم على النظرية مقارنة بنهج "الصندوق الاسود"

يمكن مقارنة تقييم الأثر القائم على النظرية مع نهج "الصندوق الاسود". هذه الأخيرة غالبا ما تعلن عن مجرد التأثير - وتهتم الدلالات الإحصائية لمعامل معالجة متوسط الأثر، ولكن لا يجري أي محاولة للإجابة على

السؤال لماذا. وقد سعت هذه الورقة إلى إظهار كيفية معالجة مسألة لماذا، والفوائد المترتبة على ذلك. ومع ذلك، فإن بعض الأمور ينبغي الانتباه لها.

لا ينبغي المبالغة في الإنتقادات الموجهة إلى التقارير التي تشير الى متوسط تأثير المعاملات، فان مسائل عدم التجانس، وكذلك مدى فهم السياق يلعبان دورا يؤدي الى تأثير معين على الأثر. ولكن نادرا أن يكون متوسط تأثير المعاملات (عادة ما يكون على حد سواء معاملة المعالجة والتي نرغب في إدخالها) ليس ذو فائدة. والواقع أنه من المحتمل جدا أن يكون هو المركز الرئيسي للاهتمام. سيكون من المضلل أن نحدد مدى معنوية الإختلاف إحصائيا، وخاصة داخل المجموعة الفرعية إذا كان متوسط تأثير التدخل عكسيا. وعلاوة على ذلك فإن التأثير المتوسط للتدخل يستخدم كأساس لحسابات فعالية التكاليف.

الثانية: يؤدي تقييم الأثر المرتكز على النظرية الى فتح السلسلة السببية بطرق مختلفة. انه يحاول فصل مختلف مراحل السلسلة السببية، ولكن أيضا إجراءات التدخل التي تؤثر وتلك التي لا تؤثر. ويمكن أن يتم هذا من خلال تحليل الإنحدار. على سبيل المثال، فإن في دراسة مشروع التغذية المتكاملة ببنجلاديش تعرض تحليل الانحدار للفجوة بين المعرفة والممارسات. لكن مثل النهج المستند الى حساب الانحدار، والذي يعتمد على نماذج إختيار العينة والمواصفات، له العديد من النقاد، الذين يفضلون إما التقييم التجريبي أو شبه التجريبي، أو التي تحدث تأثير والتي لا تحدث تأثير. ويمكن لهذه المناهج الصارمة أن تتسع لتحليل الأجزاء من عمل البرنامج. ولكن لابد من وضع التصميم للسماح بأن يختلف التدخل بإختلاف المجموعات - على سبيل المثال: بعض رجال الأعمال يحصلون على قروض، وبعض أخر يحصل على خدمات الدعم وأخرون يحصلون على الأثنين على حد سواء. عمليا، فإن تقييم الأثر المرتكز على النظرية ستجمع بين هذه التأثيرات، كما يتم مع المناهج الأخرى لتوضيح السلسلة السببية.

وأخيرا، ربما ما هو داخل الصندوق الأسود يكون فوضوي بحيث يكون في بعض الأحيان من الأفضل تركه

مغلقا؟. أختبرت دراسة للبنك الدولي أثر كهربة الريف على الخصوبة، وإن توصيل الكهرباء للمنازل يقلل إلى حد كبير من الخصوبة (البنك الدولي، 2008). وكانت هذه الدراسة قادرة على إثبات عامل واحد يمكن أن يكون على ما يبدو السبب (الحصول على المعرفة عن وسائل منع الحمل من خلال التلفزيون)، وعامل آخر لم يكن له تأثير (الحد من النشاط الجنسي). لكن هناك العديد من الوسائل الأخرى الممكنة، مثل تأثير الدخل، وغيرها من الفوائد التعليمية، وهكذا. في مثل هذه الحالات، حيث لا يمكن فصل كل التأثيرات، يكون أفضل وسيلة لقياس الأثر هو التقييم المختصر.

5. الخلاصة

تدعو هذه الورقة إلى إنتاج كمية أكبر من الدراسات الكمية الصارمة للتعرف على ما يصلح في مجال التنمية. ومع ذلك، فإن أهمية هذه الدراسات للسياسات ستكون أكبر بكثير إذا ما تم تسليط الضوء على لماذا؟، بعض التدخلات ذات أثر وأخرى ليس لها مردود. من المتفق عليه على نطاق واسع أن تقييم الأثر القائم على النظرية يمكن أن يسفر عن رؤى ضرورية. ومع ذلك، فإن العديد من الدراسات الجديدة تفشل في الوفاء بالوعد للنهج القائم على النظرية، مما يؤدي الى التكهن حول الأسباب التي أدت للتأثير، أو الى إختلافات في التأثير، بدلا من وجود تحليل تجريبي على أساس متين لشرح هذه الأسباب.

لقد قدمت مثالا لتقييم الأثر القائم على النظرية في الممارسة العملية، وكيف أن هذا النهج يؤدي مباشرة إلى الإستنتاجات التي تخدم سياسيات تعزيز أثر البرنامج. يتطلب هذا التطبيق ادماج مجموعة من المبادئ التي وضعت أعلاه. ويجب ان تصاغ نظرية البرنامج بطريقة مرنة، وتكون على إستعداد للتكيف مع الظروف المتغيرة في هذا المجال، وأن نكون واعين الى النظريات المتنافسة والعواقب غير المقصودة. تحتاج الصرامة إلى جمع كل من الواقعية، فضلا عن تحليل ماذا يحدث بدون التدخل، مما يعني إستخدام مزيج من الأساليب . يتم تحديد نظرية البرنامج في السياق الإجتماعي والسياسي والثقافي للتدخل، والتي ستكون وسيلة لتسليط الضوء على عدم التجانس المتوقع للمردود.

شكر

يعبر المؤلف عن شكره لماري غاردر للتعليقات على الصياغة السابقة من هذه الورقة. وينطبق هذا كالمعتاد. جميع الآراء المطروحة هنا هي آراء المؤلف ولا يمكن أن تؤخذ على أنها أراء مبادرة 3ie، أو أي من أعضائها أو الداعمين لها.

المراجع:

3ie guide for grantees (no date) '3ie impact evaluation practice: a guide for grantees', http://www.3ieimpact.org/page.php?pg=overview (accessed June 1, 2009).

Bamberger, Michael (2009) 'Strengthening the evaluation of program effectiveness through reconstructing baseline data' *Journal of Development Effectiveness* **1**(1): 37-59.

Blackman, Leonard and Stephanie Reich (2009) 'Randomized control trials: a gold standard with feet of clay?' in Stewart Donaldson, Christina Christie and Melvin Mark (eds.) *What Counts as Credible Evidence in Applied Research and Evaluation Practice?* [Thousand Oaks, California: Sage].

Carvalho, Soniya, Gil Perkins and Howard White (2004) 'Social funds: participation, social capital and sustainability' *Journal of International Development* **14** 611-625, 2002.

Carvalho, Soniya and Howard White (2004) 'Theory-based evaluation: the case of social funds' *American Journal of Evaluation* **25**(2) 141-60, 2004.

Centre for Global Development (2006) *When Will We Ever Learn?* [Washington D.C.: Centre for Global Development].

International Initiative for Impact Evaluation, 3ie (no date) '3ie Impact Evaluation Practice: a guide for grantees' http://www.3ieimpact.org/doc/3ie%20impact%20evaluation%20practice.pdf

(accessed June 1, 2009).

Leamer, E. (1983) 'Let's take the con out of econometrics', *American Economic Review*, **23**(1), 31–43.

Mosley, Paul and David Hulme (1996) *Finance Against Poverty* [London: Routledge].

Mukherjee, Chandan, Marc Wuyts and Howard White (1994) *Econometrics and Data Analysis for Developing Countries* London: Routledge.

NONIE (no date) 'NONIE statement on impact evaluation' http://www.worldbank.org/ieg/nonie/members.html (accessed June 1, 2009).

Ravallion, Martin (2009) 'Evaluating three stylized interventions', *Journal of Development Effectiveness* **1**(3).

Rogers, Patricia (2009) 'Matching impact evaluation design to the nature of the intervention and the purpose of the evaluation' *Journal of Development Effectiveness* **1**(3).

Save the Children (2003) *Thin on the Ground. Questioning the evidence behind World Bank-funded community nutrition projects in Bangladesh, Ethiopia and Uganda*. [London: Save the Children UK].

Weiss, Carol (1998) *Evaluation: methods for studying programs and policies*. Prentice Hall: New York.

Wawer M, Kigozi G, Serwadda D, et al. Trial of Male Circumcision in HIV+ Men, Rakai, Uganda: Effects in HIV+ Men and in Women Partners. 15th Conference on Retroviruses and Opportunistic Infections; 2008; Boston, MA; 2008.

White, Howard (2005) 'Comment on Contributions Regarding the Impact of the Bangladesh Integrated Nutrition Project' *Health Policy and Planning* **20**(6), 408-411.

White, Howard (2008) 'Of Probits and Participation: the use of mixed methods in quantitative impact evaluation" *IDS Bulletin*, 2008.

White, Howard (2009) 'Some reflections on current debates in impact evaluation' *3ie*

Working Paper No. 1 [New Delhi: International Initiative for Impact Evaluation].

White, Howard and Edoardo Masset (2006) 'The Bangladesh Integrated Nutrition Program: findings from an impact evaluation *Journal of International Development* **19**: 627-652, 2006.

White, Howard and Anju Vajja (2008) 'Can the World Bank Build Social Capital?: Community Participation in Social Funds in Malawi and Zambia' *Journal of Development Studies* **44**(8): 1145-1168.

White, Sarah (1992) *Arguing with the crocodile: gender and class in Bangladesh*, London: Zed.

Woolcock, Michael (2009) 'Toward a plurality of methods in project evaluation: a contextualised approach to understanding impact trajectories and efficacy' *Journal of Development Effectiveness* **1**(1): 1-14.

World Bank (2000) *Agricultural extension: the Kenya experience* [Washington D.C.: OED, World Bank].

World Bank (2002) *Social Funds: assessing effectiveness* [Washington D.C.: OED, World Bank].

World Bank (2005) *Maintaining Momentum to 2015? An impact evaluation of interventions to improve maternal and child health and nutrition in Bangladesh* [Washington D.C.: OED, World Bank].

World Bank (2006) *Repositioning Nutrition as Central to Development: a strategy for long term large-scale action* [Washington D.C.: World Bank].

World Bank (2008) *The welfare impact of rural electrification: a re-assessment of the costs and benefits* [Washington D.C.: IEG, World Bank].

Contacts

International Initiative for Impact Evaluation (3ie)
202-203 Rectangle One
D-4 Saket District Centre
New Delhi - 110017,India
Tel:+91 11 4989 4444
www.3ieimpact.org