



قياس مدى نجاح التدخلات التي تهدف لتحسين سبل عيش الشباب دليل عملي للرصد والتقييم

كيفين هيمبل
نايثان فيالا

قياس مدى نجاح التدخلات التي تهدف لتحسين سبل عيش الشباب

دليل عملي للرصد والتقييم

كيفين هيمبل
نايثن فيالا

يتوافر الدليل الكامل والملاحظات الفردية والمصادر الإضافية على شبكة الانترنت على <http://www.gpye.org/>

© ٢٠١١ البنك الدولي للإنشاء والتعمير/البنك الدولي

1818 H Street NW

Washington DC 20433

هاتف: 202-473-1000

الموقع على شبكة الانترنت: www.worldbank.org

جميع الحقوق محفوظة

قام بوضع هذا الكتاب العاملون في البنك الدولي للإنشاء والتعمير/البنك الدولي. ولا تعكس النتائج والتأويلات والاستنتاجات الواردة في هذا الكتاب بالضرورة وجهات نظر المدراء التنفيذيين للبنك الدولي أو الحكومات التي يمثلونها. لا يضمن البنك الدولي دقة البيانات الواردة في هذا العمل. كما لا تعتبر الحدود والألوان والتسميات، وغيرها من المعلومات التي تظهر على أية خريطة في هذا العمل بأي حال حكماً من جانب البنك الدولي فيما يتعلق بالوضع القانوني لأي إقليم أو تأييداً أو قبولاً لهذه الحدود.

الحقوق والتصرّيات

يمكن تكييف هذه المواد أو نسخها للأغراض غير التجارية فقط، شريطة أن تنسب للمؤلفين. يرجى إرسال نسخة من المواد المستنسخة لكيفن هيمبل (khempel@worldbank.org).

الاستشهاد المقترح:

هيمبل ، كيفن ونايثان فيالا. ٢٠١١. *Measuring Success of Youth Livelihood Interventions: A Practical Guide to Monitoring and Evaluation*. واشنطن، دي سي: الشراكة العالمية لتشغيل الشباب (GPYE).

<http://www.gpye.org/measuring-success-youth-livelihood-interventions>

الغلاف: كاثرين ويرثمان، المؤسسة الدولية للشباب.
الصورة: نايثان فيالا.

نحن نرحب بملاحظاتكم لمواصلة تحسين هذا الدليل.
للتعليقات والمقترحات، يرجى الاتصال بـ

كيفن هيمبل

khempel@worldbank.org

نايثان فيالا

nfiala@diw.de

التعلم لا يتحقق بالصدفة.

ولأننا ينبغي السعي إليه بحماس والإنكباب عليه باجتهاد.

— أيجيل آدمز

جدول المحتويات

xi	تمهيد
xiii	شكر وتقدير
xv	مقدمة
xv	الأساس المنطقي
xv	الجمهور المستهدف
xvi	الأهداف
xvii	النقاط التي يركز عليها الدليل
xvii	دراسات الحالة
xix	دراسة حالة صندوق العمل الاجتماعي لشمال أوغندا (NUSAF): المعلومات الأساسية
xix	نظرة عامة على الدليل وكيفية استخدامه
٣	بند ١. لماذا التقييم؟
٤	إدارة المشروع
٥	توليد المعرفة
٦	المساءلة
٧	المصداقية والاستدامة
٨	نتائج التقييم السلبي ليست بالضرورة سيئة
٨	النقاط الرئيسية
٩	دراسة حالة صندوق العمل الاجتماعي لشمال أوغندا (NUSAF): لماذا التقييم؟
٩	قراءات رئيسية
١١	بند ٢. مراجعة تصميم المشروع
١٢	تحليل المشكلة: هل نفهم المجموعة المستهدفة والسياق المحلي؟
١٤	استخدام تقييمات الشباب المتعدد القطاعات من أجل فهم مجموعتنا المستهدفة
١٩	التشخيص: ما هي المحددات المؤثرة على النواتج التي يحصل الشباب عليها؟
٢٣	الأهداف والتصميم: ما الذي نريد تحقيقه وكيف يمكننا تحقيقه؟
٢٣	دراسة حالة لصندوق العمل الاجتماعي لشمال أوغندا (NUSAF): مراجعة تصميم المشروع
٢٤	قراءات رئيسية
٢٧	بند ٣. إرساء نظام متابعة
٢٨	لماذا نحتاج إلى نظام متابعة؟
٢٨	تحديد المنطق وراء التدخل
٣٣	تحديد المؤشرات الرئيسية وأدوات جمع البيانات والافتراضات
٤٤	نظام المتابعة ورفع التقارير
٤٧	النقاط الرئيسية
٤٨	دراسة حالة لـ NUSAF: نظام المتابعة
٤٨	قراءات رئيسية
٥١	بند ٤. اختيار النوع الصحيح من التقييم
٥٢	ما هو الغرض من التقييم؟
٥٤	ربط أسئلة التقييم بتصميم التقييم
٥٨	هل يلاءم سياقنا التشغيلي نوع التقييم المنشود؟
٦٢	أنواع البرامج التي تبرر عادة إجراء تقييم للأثر

٦٥	النقاط الرئيسية
٦٦	دراسة حالة لـ NUSAF: تقرير إجراء تقييم الأثر من عدمه
٦٧	قراءات رئيسية

٧١. **بند ٥. إثبات أثر البرنامج**

٧٢	تحدي الإسناد
٧٢	ما هو في الأثر لا تحديداً؟
٧٣	كيف يمكننا تقدير الواقع المضاد؟
٧٤	أساليب إيجاد مجموعات مقارنة جيدة
٧٦	وقائع مضادة مزيفة
٧٩	النقاط الرئيسية
٨٠	دراسة حالة لـ NUSAF: تحديد واقع مضاد
٨١	قراءات رئيسية

٨٣. **بند ٦. تحديد أسلوب مناسب لتقييم الأثر**

٨٤	الاختيار من بين أساليب تقييم الأثر
٨٦	أسلوب ١: تصميم اليانصيب
٩١	أسلوب ٢: تصميم عشوائي ينفذ على مراحل
٩٣	أسلوب ٣: التصميم العشوائي الترويجي
٩٧	أسلوب ٤: تصميم عدم الاستمرارية
٩٩	أسلوب ٥: الاختلاف في الاختلاف
١٠٢	أسلوب ٦: المطابقة
١٠٥	الجمع بين الأساليب
١٠٩	النقاط الرئيسية
١١٠	دراسة حالة لـ (NUSAF): اختيار تصميم اليانصيب
١١٠	قراءات رئيسية

١١٣. **بند ٧. دليل خطوة بخطوة لتقييم الأثر**

١١٤	الإعداد لتقييم الأثر
١١٥	تحديد الجدول الزمني والميزانية
١١٧	تكوين فريق التقييم
١١٩	وضع خطة تقييم
١٢٤	وضع واختبار أداة المسح
١٣٠	إجراء المسح الأساسي والتحليل
١٣٣	إجراء مسح وتحليل متابعة
١٣٧	نشر النتائج
١٣٩	كشف المشكلات وحلها
١٤٢	النقاط الرئيسية
١٤٣	قراءات رئيسية
١٤٣	دراسة حالة لـ NUSAF: تنفيذ تقييم الأثر

١٤٥. **بند ٨. زيادة أهمية تقييم الأثر**

١٤٦	قياس مجموعة متنوعة من الآثار
١٤٨	استخدام مناهج الأساليب المختلطة
١٥٠	تحليل المنافع والتكاليف وتحليل فاعلية التكاليف
١٥٤	النقاط الرئيسية
١٥٥	دراسة حالة لـ NUSAF: زيادة أهمية تقييم الأثر

١٥٦	قراءات رئيسية
١٥٩	قائمة المراجع
١٦٧	مصادر إضافية
١٦٧	مؤلفات خاصة بتنمية/توظيف الشباب
١٦٨	مؤلفات خاصة بالمتابعة والتقييم
١٧٠	قواعد بيانات بحثية
١٧٠	الدوريات الأكاديمية
١٧١	قواعد بيانات تقييمات الأثر القائمة والجارية
١٧١	المدونات
١٧٢	بناء القدرات
١٧٢	المواقع الإلكترونية
١٧٣	ملحق ١. نموذج مؤشرات لتقييم الشباب
١٧٧	ملحق ٢. حلول التكاليف
١٨١	ملحق ٣. اختبارات التحقق والتزيف

المربعات

٥.....	الأدلة القائمة على توظيف الشباب.....	مربع ١-١
٨.....	فوائد إجراء تقييم الأثر.....	مربع ٢-١
١٤.....	أمثلة التقييمات التي أجريت على الشباب والسوق.....	مربع ١-٢
١٥.....	إطار عمل مايلز (MILES).....	مربع ٢-٢
٢٠.....	تحديد هدف تطوير المشروع.....	مربع ٣-٢
٥٧.....	أمثلة عن التقييم حسب النوع.....	مربع ١-٤
٥٩.....	دورة حياة برنامج واستراتيجيات التقييم الملائمة.....	مربع ٢-٤
٧٥.....	هل تطبيق العشوائية أمر أخلاقي؟.....	مربع ١-٥
٧٦.....	أمثلة مختارة لحالات عدم إمكانية تطبيق العشوائية.....	مربع ٢-٥
٧٦.....	الصدق الداخلي والخارجي.....	مربع ٣-٥
٧٩.....	أمثلة مختارة من التقييمات غير التجريبية.....	مربع ٤-٥
٨٨.....	مستويات العشوائية.....	مربع ١-٦
٩٠.....	مثال لتصميم اليانصيب.....	مربع ٢-٦
٩٣.....	مثال لتصميم عشوائي ينفذ على مراحل.....	مربع ٣-٦
٩٥.....	شروط ضرورية لينتج عن تصميم التشجيع تقديرات أثر صحيحة.....	مربع ٤-٦
٩٦.....	مثال لتصميم عشوائي ترويجي.....	مربع ٥-٦
٩٩.....	مثال لتصميم عدم الاستمرارية.....	مربع ٦-٦
١٠٢.....	مثال لأسلوب الاختلاف في الاختلاف.....	مربع ٧-٦
١٠٤.....	خطوات تطبيق أسلوب المطابقة.....	مربع ٨-٦
١٠٥.....	مثال على المطابقة.....	مربع ٩-٦
١١٩.....	موجز لخطة تقييم أثر.....	مربع ١-٧
١٢٢.....	مصادر محتملة للبيانات.....	مربع ٢-٧
١٢٤.....	العوامل المؤثرة على موثوقية البيانات أثناء إجراء المسح على الشباب.....	مربع ٣-٧
١٢٦.....	نموذج لموجز دليل المسح.....	مربع ٤-٧
١٢٧.....	نموذج استثمار مجلس المراجعة المؤسسية (IRB).....	مربع ٥-٧
١٢٨.....	مخطط التقرير الأساسي.....	مربع ٦-٧
١٣٠.....	أنواع الأسئلة الشائعة التي ستضاف إلى مسح المتابعة.....	مربع ٧-٧
١٣٢.....	مثال للإضافات على التقرير الأساسي بعد الاستبيان النهائي.....	مربع ٨-٧
١٣٤.....	التنقيب عن البيانات.....	مربع ٩-٧
١٣٥.....	مناذ النشر المختارة.....	مربع ١٠-٧
١٥٠.....	مثال على تقييم الأسلوب المختلط.....	مربع ١-٨

الأشكال

فوائد التقييم.....	شكل ١-١
دورة المشروع.....	شكل ٢-١
من الدليل إلى السياسة: برامج التحويلات النقدية المشروطة.....	شكل ٣-١
بيئات الشباب.....	شكل ١-٢
كيفية وضع أهداف المشروع.....	شكل ٢-٢
نظرية التدخل الأساسية لمشروع سبل عيش الشباب.....	شكل ١-٣
مكونات سلسلة النتائج وأمثلة على ذلك.....	شكل ٢-٣
النتائج المقصودة في مقابل النتائج غير المقصودة.....	شكل ٣-٣
من أسئلة التقييم إلى تصميم التقييم.....	شكل ١-٤
رسم توضيحي بصري لأثر البرنامج.....	شكل ١-٥
الأساليب التجريبية في مقابل الأساليب شبه التجريبية.....	شكل ٢-٥
مخاطر في مقارنة النتائج القبلية والبعدية.....	شكل ٣-٥
مخاطر في مقارنة المشاركين بغير المشاركين.....	شكل ٤-٥
شجرة القرار لاختيار أساليب تقييم الأثر.....	شكل ١-٦
خطوات تصميم اليانصيب.....	شكل ٢-٦
اختيار عينات البرامج الصغيرة والكبيرة.....	شكل ٣-٦
مجموعات المعالجة والمقارنة في التصميم المنفذ على مراحل.....	شكل ٤-٦
تقدير الأثر طبقاً لتصميم الترويج العشوائي.....	شكل ٥-٦
نموذج رسم بياني لتصميم عدم الاستمرارية.....	شكل ٦-٦
مثال على تحليل الاختلاف في الاختلاف.....	شكل ٧-٦
المطابقة الدقيقة حسب خمسة خصائص.....	شكل ٨-٦
نطاق الأهلية (مثال على ترتيب درجة الفقر).....	شكل ٩-٦
خطوات لإجراء تقييم للأثر.....	شكل ١-٧
نموذج جدول زمني لتقييم محتمل للأثر.....	شكل ٢-٧
مخطط لتقييم الأثر بمكون تصميمي شامل.....	شكل ١-٨
مقارنة المنافع والتكاليف.....	شكل ٢-٨

الجدول

١٧.....	نظرة عامة على القيود قصيرة المدى التي تواجه الشباب في سوق العمل	جدول ١-٢
٢٢.....	قائمة تدخلات قائمة على أدلة حسب القيود	جدول ٢-٢
٣٤.....	مثال لإطار منطقي لبرنامج تنظيم مشروعات في الوسط المدرسي	جدول ١-٣
٣٥.....	أمثلة على المؤشرات	جدول ٢-٣
٣٧.....	أمثلة على المؤشرات الخاصة بمشاريع سبل عيش الشباب	جدول ٣-٣
٤٠.....	نظرة عامة على أساليب جمع البيانات	جدول ٤-٣
٤٣.....	أمثلة على افتراضات واستجابات المشروع	جدول ٥-٣
٤٦.....	تصميم تقارير خاصة للجمهور	جدول ٦-٣
٤٧.....	المكونات النموذجية لميزانية متابعة	جدول ٧-٣
٥٣.....	أمثلة على أسئلة التقييم	جدول ١-٤
٥٤.....	الصلة بين معايير التقييم وأسئلة التقييم	جدول ٢-٤
٦٠.....	المهارات المطلوبة طبقاً لنوع التقييم	جدول ٣-٤
٦١.....	تقديرات التكاليف لأنواع التقييم المختلفة	جدول ٤-٤
٦٤.....	نظرة عامة على أنواع التقييم الرئيسية	جدول ٥-٤
١٠٧.....	نظرة عامة على أساليب تقييم الأثر	جدول ١-٦
١١٧.....	نموذج ميزانية تقييم الأثر	جدول ١-٧
١١٨.....	فريق تقييم الأثر ومسؤولياته	جدول ٢-٧
١٢٩.....	نظرة عامة على الاعتبارات الأخلاقية عند إجراء الأبحاث على الأطفال والشباب	جدول ٣-٧
١٤٧.....	فئات أسئلة تقييم الأثر	جدول ١-٨
١٥١.....	تقديرات فعالية التكاليف لبرامج حركة مناصرة الشباب Jóvenes	جدول ٢-٨
١٥٣.....	القيمة الحالية وصافي القيمة الحالية لعائد سنوي قدره ١٠٠ دولار أمريكي	جدول ٣-٨

كما

لوحظ في تقرير البنك الدولي عن التنمية في العالم لعام ٢٠٠٧، فإن المجتمع الدولي لا يزال يعلم القليل عن تأثير التدخلات الخاصة بالشباب حيث خضع عدد قليل من البرامج لعمليات تقييم نوعية. وللأسف، فهذا الأمر صحيح بالنسبة للعديد من التدخلات المتعلقة بسبل عيش الشباب في البلدان النامية. فعلى الرغم من وجود مطالب هائلة من جميع أنحاء العالم بتوفير فرص اقتصادية للشباب، إلا أننا، كمجتمع دولي، فقد كنا ولا زلنا حتى الآن غير قادرين على توفير إجابات شافية حول كيفية القيام بذلك. ومع الاعتراف بذلك، إلا أن العديد من المبادرات بقيادة المجتمع المدني والحكومات تقوم بعمل قيم للغاية. ومع ذلك، يجب علينا أن نعترف أيضاً بأن مجموع هذه التدخلات لا يصل سوى إلى جزء صغير من ١.٣ مليار نسمة من الشباب في العالم - ٩٠ في المائة منهم في البلدان النامية - ممن يستحقون الفرص والقدرات وفرص ثانية لكسب لقمة العيش الكريم لأنفسهم وعائلاتهم.

ونتيجة لعدم كفاية الأدلة، لا يحصل مدراء البرامج وصناع السياسات سوى على معلومات محدودة عن ماهية البرامج التي تحقق أكبر قدر من الفائدة للشباب المستهدفين، وأية خصائص في هذه البرامج هي الأكثر فاعلية، وما هي أفضل طريقة لتطبيقها. ولذلك، من الناحية العملية، يتم تخطيط العديد من تدخلات توظيف وسبل عيش الشباب على أساس مقابلات مع أصحاب المصالح، وسرد لقصص النجاح، ومعتقدات وخبرات مصمم البرنامج. ولكن، كي تنجح برامج سبل عيش الشباب وللسماح بتحسين وضع البرامج المستقبلية، فيجب أن تكون البرامج مستندة إلى أدلة متينة وأن يتم تنفيذها بشكل جيد وتقييمها بشكل دقيق جداً، بما يؤدي إلى توفير دروس مستفادة جديدة للممارسين وصناع السياسات وغيرهم من أصحاب المصلحة.

ولا ينعكس عدم كفاية الأدلة إلى صعوبات عند تصميم البرامج فقط، ولكنه يجعل من الصعب أيضاً الحصول على تمويل للتدخلات التي تركز على الشباب في المقام الأول. في كل مكان في العالم، يواجه صناع السياسات، بدءاً من البلدية وحتى الحكومة المركزية، طلبات ملحة في مقابل موارد محدودة. وغالباً ما تعيق ندرة الأدلة على التدخلات المتعلقة بسبل عيش الشباب تخصيص موارد عامة كافية لوضع برامج للشباب، حيث يكون أصحاب المصلحة الآخرين أكثر قدرة على الدوام على دعم القضايا التي تهمهم، مثل الاستثمارات في وسائل المواصلات أو البنية التحتية للطاقة. ونظراً لقيام صناع السياسات بتخصيص تمويلات عامة لبرامج وسياسات مدعومة بتحليلات أثر وتحليلات تكلفة ومنفعة متينة، فإنه تقع على عاتقنا مسؤولية تقديم الأدلة اللازمة على كون الاستثمار في الفرص الاقتصادية للشباب مجدياً.

ومع ذلك، غالباً ما يكون لدى المنظمات المنفذة مخاوف بشأن التكاليف والجهود اللازمة لإجراء تقييمات قوية في حين يمكنهم استخدام مواردهم الشحيحة بدلاً من ذلك لتوسيع نطاق خدماتهم بحيث تشمل مستفيدين أكثر. وللأسف، فإن التقليل من شأن الاستفادة من التقييمات يعكس قصر النظر. فعندما نفشل في التعرف بشكل منهجي على نقاط القوة ونقاط الضعف في تدخلاتنا ولا نجمع الأدلة اللازمة لإفادة البرامج الجارية والمستقبلية، فإننا نخاطر عندئذ بالحكم على مجال سبل عيش الشباب بالكامل بنقص مزمن في التمويل العام.

وإزاء هذه الخلفية، ومن أجل بناء ونشر الأدلة على أساليب مثبتة لتحسين مخرجات سبل عيش الشباب، أنشأ البنك الدولي الشراكة العالمية لتوظيف الشباب في عام ٢٠٠٨. ويشمل أعضاء الشراكة المعهد العربي لإنماء المدن والمؤسسة الدولية للشباب ومشروع فهم عمل الأطفال وشبكة توظيف الشباب. واقتناعاً بأن وجود أدلة أكثر وأفضل من شأنه تحسين تصميم ومخرجات برامج الشباب في المستقبل وتعزيز المجال بأكمله، تعمل الشراكة على تعزيز النشاط لتقييمات الأثر والمنتجات المعرفية ذات الصلة.

وقد قمنا بوضع الدليل الحالي في ظل هذا السياق. وهو يقدم عرضاً شاملاً لموضوع المتابعة والتقييم وتطبيقه العملي في مجال سبل عيش الشباب. ونأمل أن يساعدكم ويساعد منظماتكم في تحديد نوع التقييم المناسب لكم في ضوء خصائص وسياق التدخل. فإذا قررتم أن تقييم الأثر مناسب لبرنامجكم، فإن الدليل يشرح الأدوات المتاحة وخطوات العمل المحددة والتحديات العملية التي ستساعدكم في تخطيط التقييم.

ويحدونا الأمل الصادق في أن هذا الدليل سوف يساعد العاملين في المجال في اتخاذ قرارات واعية في اختيار أطر التقييم التي تفيد برامجهم ومنظماتهم وأن يساهم في قطاع سبل عيش الشباب عموماً. فإذا تمكنا مستقبلاً من استخلاص أدلة أكثر متانة من كل هذا العمل الجيد والخبرة، فسيكون لدينا صوت أقوى لإقناع صناع السياسات بتبني منهجيات ناجحة على نطاق أوسع. ونحن نتطلع إلى استمرار العمل الجماعي لصناع السياسات والعاملين في مجال التنمية وغيرهم من أصحاب المصلحة كي يقدموا لراشدي الغد الفرص الاقتصادية التي يستحقونها.

أروب بانرجي

مدير، إدارة الحماية الاجتماعية والعمل
البنك الدولي

ويليام إس. ريس

رئيس

المؤسسة الدولية للشباب

ماركوس بيلغريم

مدير

شبكة توظيف الشباب

فوريو روزاتي

مدير

مشروع فهم عمالة الأطفال

تشرين الثاني/نوفمبر ٢٠١١

شكرو تقدير

لم يكن إعداد هذا الدليل ممكناً لولا مساهمات العديد من الزملاء والأصدقاء. أولاً وقبل كل شيء، نود أن نشكر ماتياس لندبرغ وويندي كينغهام لتوجيههما العام وملاحظتهما ودعمهما طوال عملية التخطيط والكتابة، وكذلك مدير وحدة توفير الحماية الاجتماعية والعمل بالبنك الدولي، أروب بانيرغي.

كذلك، نود أن نعبر عن شكرنا وتقديرنا لزملائنا من الشراكة العالمية لتوظيف الشباب (GPYE) الذين قدموا مساهمات كبيرة في توفير المحتوى والمقترحات لأجزاء هذا الدليل. ويشملون إبراهيم التركي من المعهد العربي لإنماء المدن؛ عويس صوفي وترفيس آدكنز ودان أوليفر وجاك بيزوللو وهنا كوري من المؤسسة الدولية للشباب؛ وماركوس بيلجرم وسوزانا غوانزاليس بورتو ودرو غاردنر من شبكة توظيف الشباب؛ وغابريلا بريجليا من مشروع فهم عمل الأطفال. حيث كانت كتابة هذا الدليل حقاً جهداً جماعياً.

كما أننا ممتنون للغاية لزملائنا المراجعين غلوريا لا كافا من البنك الدولي وفيونا ماكولاي وفيرونیکا توريس من مؤسسة ميكنغ سنتس الدولية.

كما نود أن نعبر عن تقديرنا للعديد من زملائنا داخل وخارج البنك الدولي الذين تحملوا عناء قراءة المسودات السابقة لهذا الدليل وتقديم ملاحظات ومقترحات ومواد أساسية لا تقدر بثمن. ويضم مساهمو البنك الدولي:

بالوما أسيفيدو وجوليانا بيلزرا ايز وكارلوس أسينجو رويز وستيفاني برودمان وشوبها تشاكرافاتي ووينونغ تشو وفولوا اردمانسدويرفير وتانغا لوهمان وفلورينتيا مولاغ وسليمان نمارا وديفيد لوك نيهواوس وأزرا كاكابور نوريك وبا أوبونغ وميشيل ريبوسيو وليوبولد سار وحنين إسماعيل سيد وديفن سيلفر وستافروس جورج وكورنيليا تيسليوك.

وكذلك نتوجه بالشكر لدينو ليناريز من كوليكتيفو انتيغرا دي ديسارولو Colectivo Integral de Desarrollo وبيا سوندرز من مؤسسة التعليم من أجل التوظيف؛ وألكساندر كوليف من مركز التدريب الدولي لمنظمة العمل الدولية (ILO-ITC)؛ وديفيد روساس من مصرف التنمية للبلدان الأمريكية؛ وكريستين هوسوتر وبيتينا سيليرناغل من الجمعية الألمانية للتعاون الدولي (GIZ)؛ وأن غولا من المركز الدولي للبحوث المتعلقة بالمرأة؛ وكارولين جينر وشانون ويندت من منظمة إنجاز الشباية؛ وويتني هارلسون من مؤسسة ميكنغ سنتس الدولية؛ وريوا ميسرا من مؤسسة ماستر كارد؛ وليا كاتيرينغ وسكوت روديك من جمعية مينونايت للتنمية الاقتصادية (MEDA)؛

وسونيا سيلفا وديفيد وولكومب من المؤسسة الدولية لطفل السلام؛ وكارين أوستريان من مجلس السكان؛ وراني ديشباندي من مؤسسة إنقاذ الأطفال؛ وجاستن سايكس من مبادرة صلتك (Silatech)؛

ونادر قباني من الأمانة السورية للتنمية؛ ورايتشل بلم من الوكالة الأمريكية للتنمية (USAID)؛ وهيلين غايل من منظمة شباب الأعمال الدولية.

بالإضافة إلى ذلك، نود أن نعبر عن شكرنا وتقديرنا لبول غيرتير وسيباستيان مارتينيز وباتريك بريماند ورولا رولينغز وكريستال فيرميرش على مؤلفهم الممتاز *Impact Evaluation in Practice*. حيث يبنى هذا الدليل على عملهم مع تكييف بعض المواد والرسوم التوضيحية بحيث تلائم ميدان توفير سبل العيش للشباب وتقديم عرض أكثر إيجازاً لأساليب تقييم الأثر. كذلك، نود أن نشكر المشاركين في ورشات عمل لتقييم شبكة توظيف الشباب في نيروبي وبيروت وجنيف والعديد من الأعضاء النشطين في موقع مجموعة YEN؛ والمشاركين في المؤتمر الشبابي العالمي لمؤسسة ميكنغ سنتس الدولية والمؤتمر العالمي للشباب ٢٠١٠، والذين شكلت جميع أسئلتهم واستفساراتهم محتوى هذه الوثيقة.

شكر خاص نوجهه للمدقق كريس روسش، لما قدمه من مشورة قيمة واستجابة ورغبة في العمل في ظل إطار زمني ضيق جداً. كما نشكر أيضاً كارولين اسكلايز وغيليان ماك كاليون وكاثرين ويرثمان لمساهماتهم في تصميم هذا العمل.

أخيراً، يود كيفن هيمبل أن يشكر كارلا روغاس على رحابة صدرها.

أية أخطاء قد تظل في هذه الوثيقة هي أخطاء خاصة بنا.

مقدمة

تنفذ

البرامج التي تهدف إلى توفير الدعم الفعال لآفاق توظيف الشباب منذ قرون في البلدان الصناعية، ولكن تعد هذه البرامج جديدة نوعاً ما في البلدان النامية. وبمعنى أعم، توفر التدخلات المتعلقة بسبل عيش الشباب دعم الأساليب المتاحة للشباب من أجل كسب قوتهم، وتشمل هذه التدخلات التدريب والخدمة العامة وتنظيم المشروعات الشبابية والخدمات المالية. وبشكل أكثر تحديداً، يعرف العديد من العاملين في المجال برامج سبل عيش الشباب على أنها نشاطات تستهدف بشكل خاص الجماعات المستضعفة والمهمشة في الاقتصاد غير الرسمي، مع التركيز بوجه خاص على العمل الحر. يتبنى هذا الدليل التعريف الأعم ويشمل تنمية القوى العاملة للقطاع الرسمي.

ونظراً لكون هذا المجال جديداً ومبتكراً نسبياً، فقد تم تقييم عدد قليل من التدخلات بشكل حثيث. ففي الحقيقة، لم يستشهد غالبية العاملين في المجال سوى بمقدار قليل من الأمثلة. ولكن ما الذي تعنيه حقاً كلمة حثيث؟ ما هي الأساليب التي تكون حثيثة بشكل كاف، وما هي الأساليب التي لا تكون حثيثة؟ بالنسبة للعاملين في المجال، عادة ما يكون واضحاً أن التدخل الذي ننفذه يعطي النتائج المطلوبة، فلماذا ننفق مواردنا المحدودة على عملية تقييم باهظة الثمن إذا كان بمقدورنا استخدام المال لتقديم الخدمات لمزيد من الشباب بدلاً من إنفاقها على التقييم؟

أما بالنسبة للأشخاص الذين لا يشاركون بشكل مباشر في التدخل، فعادة لا تكون فاعلية التقييم واضحة. حيث يحتاج صناع القرارات والمناحون إلى نتائج موثوقة وذات شفافية تلبى بعض المعايير الدنيا للموثوقية. وعادة ما يبحثون عن تقييمات تستخدم أساليب البحث المثبتة في مجال العلوم الاجتماعية، والتي يمكن أن تقدم تقديرات متينة حول كيفية تأثير تدخل ما على المشارك النموذجي في البرنامج. وعلى الرغم من اهتمام العاملين في المجال بتوفير المعلومات ذات جودة حول برامجهم، إلا أنهم قد يشعرون بأن عمليات التقييم الحثيثة، بتعقيدها وتكاليفها المحتملة وما تتطلبه من الموارد الأخرى، غالباً ما تكون غير واقعية وبعيدة المنال.

الجمهور المستهدف

يعد هذا الدليل دليلاً تمهيدياً تمت كتابته للعاملين في المجال الذين ليست لديهم معرفة – أو ممن لديهم معرفة محدودة للغاية – حول تقييم الأثر أو أساليب البحث الكمي، ولكنهم مع ذلك يولون اهتماماً بإثبات النتائج الحقيقية لعملهم. يخاطب هذا الدليل مدراء البرامج وموظفي المتابعة والتقييم المحليين في كافة أنماط المنظمات الناشطة في مجال سبل عيش الشباب: المؤسسات غير الحكومية المحلية والدولية، والموظفين الحكوميين المحليين والدوليين، والجهات المانحة الثنائية والمتعددة الأطراف.

ونظراً لتنوع الخلافات والخبرات بين العاملين في المجال، يستحيل أن يتم تصميم هذا الدليل بحيث يلي احتياجات الجميع بشكل متساو. إلا أننا قد حاولنا تقديم نقاش شامل حول أساليب تقييم التدخلات المتعلقة بسبل عيش الشباب حتى يتمكن القراء من تحديد الأقسام الأكثر ملاءمة لاهتماماتهم واحتياجاتهم.

الأهداف

نهدف بتقديم هذا الدليل إلى تزويد القراء بالمجموعة الأساسية من المبادئ والأدوات المطلوبة لاتخاذ قرارات مدروسة بشأن كيفية تقييم برامجهم بشكل أفضل. حيث نسعى إلى تقديم فهم واضح لخيارات التقييم المتنوعة المتاحة وللاعتبارات التي ستسمح للعاملين في المجال بانتقاء الخيار الأنسب استناداً إلى أهداف التعليمية والسياق التشغيلي. وفضلاً عن ذلك، نقدم وصفاً لكيفية إدارة تقييم الأثر إذا كان تقييم الأثر هو أسلوب التقييم الذي تم اختياره.

وتمثل هدفنا الأسمى في تقوية أساس عملية وضع البرامج وصنع القرارات بشكل سليم من خلال زيادة عدد عمليات تقييم النوعية في مجال سبل عيش الشباب، وبالتالي تيسير توسيع نطاق التدخلات وتكرار النجاح منها.

.....

- مجالات التدخل لبرامج تنمية سبل عيش الشباب
- التدريب وتنمية المهارات.
- التوظيف المدعوم، بما في ذلك دعم الأجور والأشغال العامة وبرامج الخدمة العامة.
- خدمات التوظيف، بما في ذلك المساعدة في البحث عن الوظائف ودعم التنسيب.
- المؤسسات والمشروعات الشبابية.
- الخدمات المالية المقتصرة على الشباب.
- برامج غير تقليدية للمجموعات المهمشة.
- أحكام سوق العمل التي تؤثر على الشباب.

المصادر: بيتشرمان وآخرون (٢٠٠٧)؛

كننغهام وسانشينز بويرتا وويرملي (٢٠١٠)؛

DFID (١٩٩٩).

النقاط التي يركز عليها الدليل

يتناول الدليل عمليات متابعة وتقييم التدخلات المتعلقة بسبل عيش الشباب، مع التركيز بشكل خاص على تقييم الأثر. عادة ما يقترن استخدام مصطلحي المتابعة والتقييم، إلا أنهما يشيران إلى أنشطة تختلف عن بعضها.

يتبع التقييم تنفيذ وتقديم سيرة عملية تدخل ما من أجل تقديم الدعم لإدارة البرنامج. والمتابعة:

- تتضمن جمع البيانات الخاصة بتنفيذ تدخل ما ومؤشرات النتائج.
- تحديد الامتثال لخطط العمل والميزانيات.
- تستخدم المعلومات لإدارة المشروع واتخاذ القرارات.
- تكون مستمرة.
- تجيب عن السؤال «هل نقوم بتنفيذ المشروع بالشكل الصحيح؟»

التقييم يتم فيه تقييم تصميم أو تنفيذ أو نتائج تدخل ما من أجل دعم عملية التخطيط الجديدة. والتقييم:

- يتضمن جمع البيانات الخاصة بتصميم أو تنفيذ أو نتائج مشروع ما.
- يأخذ بعين الاعتبار صلة مشروع ما بالحاجة المطلوب تلبيتها وكفاءة وفعالية هذا المشروع واستدامته.
- ينتج معلومات مفيدة حول أثر التدخل.
- يكون دورياً، وعادة ما يتم إجراؤه سنوياً عند انتهاء مشروع ما، ويتضمن المتابعة.
- يجيب عن السؤال «هل نقوم بتنفيذ المشروع الصحيح؟»

من الناحية المثالية، يجب أن تكون عمليتي المتابعة والتقييم جزءاً لا يتجزأ من أي برنامج ويجب التخطيط لهما في مرحلة تصميم البرنامج. في الحقيقة، قد لا يكون من الممكن تقييم نجاح تدخل ما بشكل دقيق إذا ظل التقييم فكرة مستدركة تعطى أولوية ضئيلة حتى ينتهي المشروع.

تقييم الأثر هو أحد أنواع التقييم يتم فيها قياس التغيرات في رفاهية الأفراد والعائلات أو التجمعات المنسوبة إلى تدخل معين. يجيب تقييم الأثر عن السؤال: ما الذي كان يمكن أن يحدث للمستفيدين إذا لم ينفذ البرنامج؟ على سبيل المثال، إذا عثر خريج جديد من برنامج التدريب على المهارات على وظيفة، فهل هذا يعتبر محصلة مباشرة للبرنامج، أو هل كان لذاك الشخص أن يعثر على عمل بأية حال؟ بمقارنة النواتج التي يلمسها المشاركون بالنواتج التي تلمسها مجموعة من غير المشاركين الذين تم اختيارهم بشكل جيد لإجراء المقارنة معهم، فإنه يصبح من الممكن فهم السببية. وبكلمات أخرى، تمكنا عمليات تقييم الأثر من أن نغزو أية تغيرات ملحوظة في رفاهية المستفيدين من برنامج ما إلى فاعلية التدخل الذي قمنا به.

يعد تقييم الأثر واحداً من بين العديد من أنواع التقييم المتاحة، بمزاياه وأوجه القصور فيه. ونحن نعتقد أن تقييم الأثر غير مطلوب لكل تدخل وأن التقييم يجب أن يدعم عملية وضع البرامج، وليس العكس. ينبغي أن تكون أية عملية تقييم مناسبة للخصائص التشغيلية للتدخل المعني وأن تكون مناسبة أيضاً لسياق هذا التدخل، بحيث يتم دمج التقييم في إطار أوسع يبنى على نظام متابعة تم إرساؤه. وهذا يعني أننا نعتقد بأنه من الممكن تعلم المزيد عند استخدام منهجيات تقييم الأثر بشكل أكثر تواتراً.

[تعريف]

المتابعة: عملية مستمرة يتم فيها جميع وتحليل المعلومات لرؤية إلى أي مدى يتم تنفيذ مشروع ما أو برنامج ما أو سياسة ما بشكل جيد ومدى مطابقة الأداء للنتائج المتوقعة.

التقييم: عبارة عن تقييم منهجي وموضوعي لتصميم وتنفيذ ونتائج مشروع جاري تنفيذه أو تم تنفيذه من أجل تحديد صلة هذا البرنامج بالحاجة المراد تلبيتها ومدى تحقيق الأهداف والكفاءة والفاعلية والأثر والاستدامة.

تقييم الأثر: نوع خاص من التقييم يعمل على تقييم التغيرات في رفاهية الأفراد والأسر والتجمعات التي يمكن أن تترى إلى تدخل معين.

المصدر: تم الاقتباس من غيرتزر وآخرون (٢٠١١)، كوزيك وريست (٢٠٠٤): OECD (١٩٩١).

- يختلف هذا الدليل عن الأعمال الأخرى بثلاث نواحي عامة:
أولاً، نقوم مباشرة بتطبيق مبادئ المتابعة والتقييم - لا سيما مبادئ تقييم الأثر - على قطاع سبل عيش الشباب. يقدم الكتاب أمثلة من واقع الحياة وشهادات ومؤشرات وتحديات عملية ذات صلة بتقييم التدخلات المتعلقة بسبل عيش الشباب.
- ثانياً، نسعى إلى تحقيق توازن بين حقبة الأدوات العملية التي تركز على المتابعة والتقييم بشكل عام (مثال، غوسباريني وآخرون ٢٠٠٤؛ كيلوغ ١٩٩٨؛ كيلوغ وآخرون ٢٠٠٤) وغيرها من المطبوعات التي تركز بشكل خاص على تقييم الأثر (مثال، بيكر ٢٠٠٠؛ دوفلو وغلينرستر وكريمر ٢٠٠٦؛ غيرتير وآخرون ٢٠١١؛ غاندر وكولوال وصمد ٢٠١٠؛ رافاليون ٢٠٠٨).
- ثالثاً، نستهدف بشكل صريح العاملين في مجال سبل عيش الشباب الذين لا يكون لديهم معرفة مسبقة بأساليب البحث والتقييم والذين يطلبون توضيحاً مقتضباً، ولكن شامل، للمتابعة والتقييم وكيفية تطبيقها في عملهم اليومي. لذلك، على النقيض من المطبوعات أعلاه، تم تصميم هذا الدليل لتقديم محتوى محدد ومتمحور حول الشباب. للإطلاع على مقدمة أشمل في الموضوع الخاص بتقييم الأثر وممارسته في مجال التنمية، فإننا نشجع القارئ على الرجوع إلى *Impact Evaluation in Practice* الذي أعده غيرتير وزملاؤه (٢٠١١).

دراسات الحالة

نستخدم في الأقسام المختلفة للدليل مشروع صندوق العمل الاجتماعي لشمال أوغندا (NUSAF) لتوضيح النقاط الرئيسية في كل بند. لقد اخترنا (NUSAF) لأغراض هذا الدليل لأنه يعكس واجهات كثيرة من جوانب برنامج معياري لسبل عيش الشباب ولأن عملية تقييم الأثر لهذا المشروع تعاملت مع العديد من التحديات. وباعتراف الجميع، يعتبر مشروع (NUSAF) كبيراً بالمقارنة مع العديد من المشروعات الأخرى المتعلقة بسبل عيش الشباب. ولكن كما سنرى، تكون عمليات تقييم الأثر ممكنة كذلك للبرامج الأصغر. نأمل أن يجد القراء أن جوانب دراسة الحالة قريبة بشكل كافٍ من وضعهم.

دراسة حالة صندوق العمل الاجتماعي لشمال أوغندا (NUSAF): المعلومات الأساسية

معلومات عامة	
اسم المشروع:	برنامج فرص الشباب التابع لصندوق العمل الاجتماعي لشمال أوغندا «NUSAF»
المجموعة المستهدفة:	الشباب الفقراء بين سن ١٥-٣٥ سنة في منطقة ما بعد الصراع في شمال أوغندا
عدد المستفيدين:	٨,٠٠٠ +
الميزانية:	١.٦ مليون دولار أمريكي
(...)	

(...) دراسة حالة صندوق العمل الاجتماعي لشمال أوغندا (NUSAF): المعلومات الأساسية



سياق المشروع

شهدت أوغندا طوال عقدين نمواً اقتصادياً وأمنياً مادياً واستقراراً سياسياً، بالإضافة إلى ارتفاع في مستويات التعليم والصحة. غير أن المنطقة الشمالية ظلت متأخرة عن باقي البلاد من كل النواحي. فطالما تركزت الحركة التجارية منذ القدم في جنوب غرب ووسط أوغندا وذلك يعود لأنماط التطور في حقبة ما قبل الاستعمار وخلالها، والقرب من الشركاء التجاريين وتوفير البنية التحتية.

إضافة إلى ذلك، زعزعت الحرب الأهلية التي دامت عقدين بالإضافة إلى عدم الأمان في الشمال (وفي الدول المجاورة) اقتصاد المنطقة ومجتمعها. فقد شهدت كل المناطق في الشمال تقريباً شكلاً من أشكال عدم الأمان المادي - من تمرد مسلح ونزوح داخلي وسرقة الماشية وغيرها. حديثاً، انتهت حرب أهلية في مقاطعات آتشولي الأصولية كانت قد تسببت في نزوح كل سكان الريف، أي ما يقارب مليوني نسمة. وفيما تضاءلت حالات الطوارئ الإنسانية، خفت المساعدات الإنسانية تدريجياً وازدادت المساعدات التنموية الوطنية والدولية بشكل كبير.

كانت تهدف خطة حكومة أوغندا للسلام والانتعاش والتنمية إلى تعزيز سلطة الدولة وإعادة بناء المجتمعات والترويج للسلام والمصالحة وإعادة إحياء الاقتصاد من خلال باقية من برامج عدة، وكان NUSAF أحد هذه البرامج.

أنشطة المشروع

استهدف برنامج توفير الفرص للشباب المتعلق بـ NUSAF شباب تتراوح أعمارهم بين ١٥ و ٣٥ عاماً والذين عاشوا في ظل ظروف الفقر وعانوا من البطالة ونقص فرص التوظيف. وهكذا نظم الشباب أنفسهم في مجموعات صغيرة وحددوا مهارة مهنية تهتمهم ومعهد للتدريب المهني يودون الانتساب إليه وقدموا طلباً أمام المكاتب التقنية لـ NUSAF في المقاطعة للحصول على التمويل.

تألف برنامج الفرص للشباب من مكونين أساسيين.

وفر المكون الأول تحويلاً نقدياً يناهز ٧,٠٠٠ دولار لمجموعات شباب محلية تستخدم فيها المجموعات هذه الأموال للالتحاق بمعاهد التدريب المهني وشراء مواد التدريب ودفع تكاليف إطلاق المشروع ليتمكنوا من ممارسة هذه الحرفة بعد التخرج.

أما المكون الثاني، فبنى قدرة المنظمات غير الحكومية ومنظمات المجتمع المحلي ومعاهد التدريب المهني لتستجيب لاحتياجات الشباب. (فقد دمرت طول فترة النزاع وحدته الكثير من البنية التحتية في شمال أوغندا، وخصوصاً مؤسسات التعليم. لذلك كان الاستثمار في هذه المؤسسات السبيل لبناء القدرات المستقبلية).

المصدر: ياتلمان ومارتينز وفيالا (٢٠٠٩).

نظرة عامة على الدليل وكيفية استخدامه

يقدم الدليل على شكل سلسلة من البنود القصيرة متشعبة من قسمين أساسيين. يتعلق القسم الأول بفهم أسباب التقييم والتحضير له، في حين يتعلق القسم الثاني بالتنظيم لتقييم الأثر. ومع أنه من المهم الاطلاع على كافة نواحي العملية، فليس من الضروري قراءة الدليل من المقدمة إلى الخاتمة. بدلاً من ذلك، تمت صياغة كل بند ليكون فصلاً مستقلاً عن غيره يمكن قراءته بمعزل عن باقي الفصول، بحسب احتياجات كل قارئ. إذا كنت تود الاطلاع أكثر على التخطيط للمتابعة والتقييم بشكل عام، ابدأ بالقسم الأول. أما إذا كنت على اطلاع بالمتابعة والتقييم وتود معرفة المزيد عن تقييم الأثر، فستجد القسم الثاني أقرب إلى ما تبحث عنه. يشير دليل القارئ التالي إلى البنود التي تعكس حاجة كل فئة من القراء.

دليل القارئ

القسم الأول: وضع أسس لعملية تقييم تصف البنود الأربعة في هذا القسم كيفية التحضير للتقييم.					
البند	الوصف	صناع السياسات	مدراء المشروعات	مسؤولو المتابعة والتقييم	موظفو البحث والسياسات
١	يناقش سبب أهمية التقييم وكيف يقوم بدعم وضع البرامج والأهداف التنظيمية.	✓	✓		
٢	يراجع بعض الأسئلة الهامة حول تصميم البرنامج والتي يجب الإجابة عنها قبل الانتقال إلى المتابعة والتقييم.		✓		
٣	يقدم الخطوات الأساسية في تطوير نظام المتابعة الذي يعتبر الأساس لأي تقييم.		✓	✓	
٤	يستفسر عن نوع تقييم الذي يناسب برنامج محدد. تختلف الإجابة بحسب الأهداف التعليمية وسياق المشروع وخصائصه والموارد المتوفرة.	✓	✓	✓	✓
القسم الثاني: تعزيز عملية التعلم من خلال تقييم الأثر تقدم البنود الموجزة في هذا القسم تقييم الأثر وتوفر إرشاداً واضحاً حول تنفيذه في سياق البرامج المتعلقة بسبل عيش الشباب.					
البند	الوصف	صناع السياسات	مدراء المشروعات	مسؤولو المتابعة والتقييم	موظفو البحث والسياسات
٥	يقدم أبرز مميزات تقييم الأثر ويفسر لماذا لا تستوفي بعض أساليب التقييم المستخدمة عادة معايير الجودة عينها.	✓	✓	✓	✓
٦	يراجع أدوات وأساليب إجراء تقييم الأثر ويفسر كيف تعمل وماذا تتطلب. كما يعرض شجرة قرار لمساعدة القراء على التفكير في الطريقة المثلى التي تناسب حالتهم الخاصة.		✓	✓	✓
٧	ينتقل من المستوى النظري إلى المستوى العملي بوصف أبرز الخطوات التي تدخل في تنفيذ تقييم الأثر وتوفير موارد عملية. تغطي هذه الخطوات كامل العملية، من التحضيرات الأولية إلى نشر النتائج.		✓	✓	✓
٨	يقدم أدوات لزيادة أهمية تقييمات الأثر. كما يتضمن لمحة عامة عن مجموعة متنوعة من أسئلة تقييم الأثر وعن استخدام طرق مختلطة بالإضافة إلى تحاليل التكلفة الفاعلة وتحاليل التكلفة والمنفعة.				✓

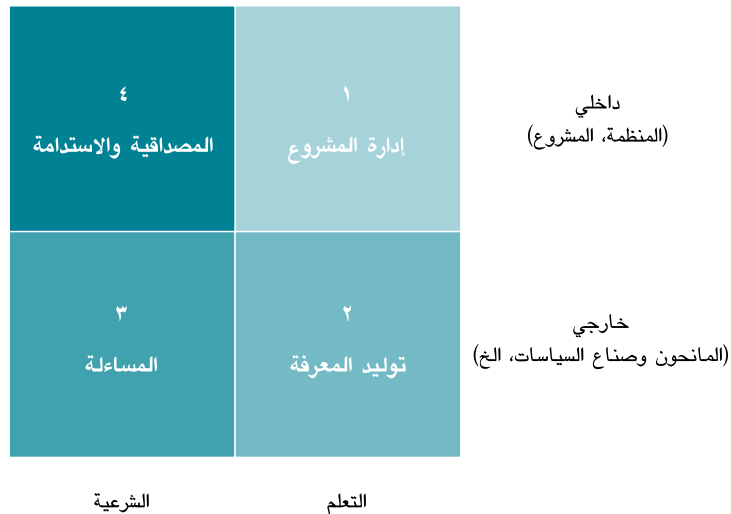
النجاح يتوقف على معرفة الأمر المفيد.

— بيل غيتس

إن التدخلات التي تركز على الشباب معقدة بطبيعتها. لأننا نتعامل مع فئة مستهدفة ديناميكية تمر في مرحلة انتقالية سواء بيولوجياً أو اجتماعياً أو قانونياً، فالتدخلات التي نقوم بها تكون في غاية التنوع بطبيعتها وتمتد نتائجها على مجموعة من القطاعات. لذا فتقييم هذه التدخلات بالطريقة المناسبة، ولو كان يشكل تحدياً، يعد مكوناً رئيسياً في وصفة النجاح. يتمثل الهدف من هذا البند في تقديم لمحة عامة عن كيفية استفادة كل منظمة على حدة والمجال ككل من المعرفة التي تحصل عليها من التقييم الرسمي، وبالأخص من خلال تقييم الأثر. ونحن نرى أنه ثمة دافعين رئيسين للتقييم: التعلم وإرساء الشرعية. ويوجد لكل دافع جمهور داخلي وخارجي (أنظر شكل ١-١). ينتج عن هذين الدافعين معاً أربعة أسباب جيدة لإجراء التقييم:

- إدارة المشروعات
 - توليد المعرفة
 - ضمان المساءلة
 - تعزيز مصداقية واستدامة المنظمة
- تتم في ما يلي مناقشة هذه الأسباب.

شكل ١-١ فوائد التقييم



إدارة المشروع

تسمح التقييمات برؤية قيمة عملنا الفعلية. فمعظمنا يريد معرفة الفرق الذي تحدثه برامجنا في حياة الشباب الذين نخدمهم. هل حقق مشروعا النتائج المرجوة؟ من استفاد أكثر، ومن كانت استفادته أقل؟ تساعد التقييمات في الإجابة عن هذه الأسئلة وغيرها عبر تقييم الحاجة للتدخل المعني وفاعليته وكفاءته وأثره واستدامته.

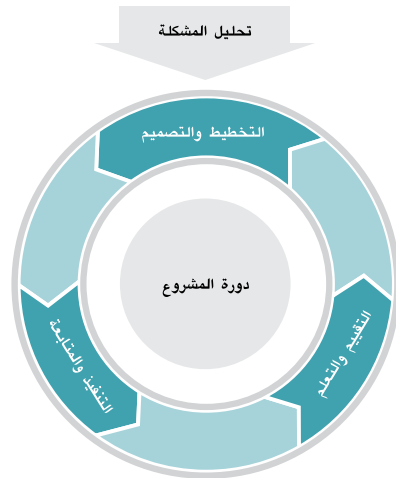
عمليات التقييم تعزز التعلم. بتقييم تصميم التدخل وتنفيذه ونتائجه، تعزز التقييمات التعلم في المنظمة. حيث يسمح لنا بتحديد أي من أجزاء تدخلنا كانت ناجحة وأي منها لم ينجح كما كان مخططاً. فيمكن مثلاً لتقييم ما أن يكشف أن كم التدريب الذي تم تقديمه من خلال تدخل ما لم يكن كافية، لما نتج عنه من نسبة تعلم منخفضة، أو على العكس أن التعلم كان مكثفاً جداً وقد أرهاق الطلاب فأدى إلى الانسحاب. وبنفس الكيفية، يمكن للتقييم أن يساعدنا على فهم نتائج مشروعنا التي لم نكن نقصد حدوثها، مثل زيادة نسبة استهلاك أولياء الأمور للكحول نتيجة توفير فرص مدرة للدخل للفتيات.

تدعم عمليات التقييم التخطيط الجديد. توفر التقييمات لنا، نحن مدراء البرامج، المعلومات التي نحتاج إليها لاتخاذ قرارات استراتيجية حول التغييرات الضرورية التي نحتاج لإجرائها على تصميم المشروع أو التخطيط له أو تنفيذه. فعلى الرغم من أن التقييمات بشكل عام (وتقييمات الأثر بشكل خاص) تنتج المعلومات على نحو دوري وليس بشكل مستمر، فهي تشكل مع ذلك أجزاء قيمة من دورة المشروع. حتى التقييمات التي تتم بأثر رجعي تتطلع بشكل أساسي إلى الأمام فيما يتعلق بعجل البرامج اللاحقة (اليونيسيف ١٩٩١). كما يظهر في شكل ٢-١، يطبق التقييم الدروس الناتجة من تدخلات جارية أو منتهية على تخطيط البرامج الحالية والمستقبلية وتصميمها. ويساعد التقييم إذا صمم جيداً أصحاب العلاقة على القيام بالتعديلات الضرورية لتقليص تمويل برامج الشباب التي لا تحقق أهدافها، فيما يبقون على البرامج الأخرى التي تحقق أهداف جيدة أو يمكنها تحقيقها. فبدون الحصول على بيانات من تقييمات جيدة، يصبح خطر التوصل إلى استنتاجات خاطئة حيال أي البرامج يتوجب استمرارها وكيفية توزيع الموارد أكثر وقعاً. (البنك الدولي ٢٠٠٩).

«لقد تمكنا بحصولنا على بيانات ملائمة حول سبب انسحاب الشباب من التدريب من تقرير منح مخصصات. وبذلك انخفض معدل الانسحاب من ٣٥ بالمائة إلى ٩ بالمائة».

— مدير البرنامج،
منظمة كاريبية غير حكومية

شكل ٢-١ دورة المشروع



توليد المعرفة

يتصف مجال سبل عيش الشباب بنقص شديد في الأدلة القوية. حتى لو كانت مؤسساتنا تبني جمع البيانات بانتظام لمتابعة وإجراء تقييمات الأداء المعيارية (كما هو الحال من خلال القيام بمقارنات بسيطة بين قبلية وبعدية للمشروع والمجموعات البؤرية/المركزة)، غالباً ما نفشل في بناء معرفة يمكن تعميمها لإفادة المجال بأكمله (سيفدوف وليفين ويردسال ٢٠٠٦). تتطلب حيازة هذه المعرفة عادةً دراسات للأثر تستخدم منهجيات محددة لتوفير تقديرات موثوقة لنجاح تدخل معين. بالرغم من إنفاق مليارات الدولارات على تنفيذ برامج سبل عيش الشباب، تتوفر دراسات أثر قليلة نسبياً. فمثلاً، في مراجعة عالمية لنتائج التدخلات لتوظيف الشباب، وجد بيتشرمان وزملاؤه (٢٠٠٧) ثلاث تقييمات فقط ذات جودة لبرامج تنظيم مشروعات الشباب. على نحو مماثل، لا نعلم الكثير عن الاستراتيجيات الأخرى للترويج لسبل العيش، كتوفير فرصة ثانية في التعليم أو وضع برامج الأشغال العامة أو تنمية المعرفة والخدمات المالية للشباب (أنظر مربع ١-١). ومع أن الجهود تزايدت لبناء أدلة قوية في السنوات الأخيرة، إلا أننا ما زلنا بحاجة إلى الكثير من المعرفة.

مربع ١-١ الأدلة القائمة على توظيف الشباب

أجرى بيتشرمان وآخرون (٢٠٠٧) مراجعة عالمية للتدخلات المتعلقة بتوظيف الشباب ووجدوا أن «برنامجاً واحداً فقط من أصل عشرة يتضمن تقييماً يقيس كلا من الأثر الصافي والتكلفة». أما أنواع التدخلات التي وجد أنها تعاني من أكبر الفجوات في المعرفة، فهي برامج التوظيف المدعومة مالياً، وتنظيم مشروعات الشباب، وخدمات التوظيف والإصلاحات التنظيمية. وعلى المستوى الإقليمي، كانت الأدلة قليلة بشكل خاص في آسيا والشرق الأوسط وشرق وشمال أفريقيا والمنطقة الأفريقية جنوب الصحراء الكبرى. تتوفر قاعدة بيانات محدثة للتدخلات لتوظيف الشباب والتقييمات على الموقع الإلكتروني لسجل وظائف الشباب (<http://www.youth-employment-inventory.org>).

أجرى كارد وكولف وويبر (٢٠٠٩) تحليل ميثا (دراسة ما وراء التحليل) لبرامج سوق العمل النشطة في منظمة التعاون الاقتصادي والتنمية (OECD). وبمقارنة أنواع البرامج، وجدوا أن برامج التوظيف المدعومة من القطاع العام تكون ذات توقعات أثراً أقل إيجابية. أما برامج المساعدة في البحث عن الوظائف فتكون ذات آثار إيجابية نسبية على المدى القصير، في حين تميل برامج التدريب الصيفية والتدريب في مواقع العمل إلى إبداء نتائج أفضل على المدى المتوسط مقارنة بالمدى القصير. وقد وجد المؤلفون أن البرامج التي تستهدف الشباب في (OECD) ذات فرصة أقل في الإسفار عن أثر إيجابي مقارنة بالبرامج التي لا تستهدف فئة معينة.

يلخص إيلاربان وروساس شادي (٢٠٠٩) نتائج التقييمات بالغة الدقة لبرامج التدريب على وظيفة في أمريكا اللاتينية. خلافاً للأدلة في البلدان المتطورة، تشير النتائج إلى الآثار الإيجابية على التوظيف وجودة الوظائف للمتدربين، وبخاصة في أوساط النساء والمشاركين الأصغر سناً. ويقر التقرير بأن فجوة المعرفة لا تزال كبيرة حول الآثار طويلة المدى لمثل هذه التدخلات في أمريكا اللاتينية.

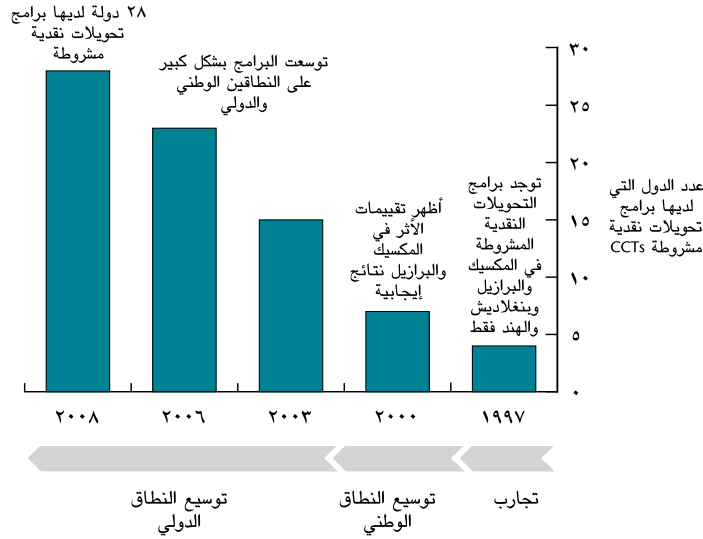
يلخص كينغهام وسانشين-بويرتا وويرملي (٢٠١٠) بيان الأدلة المتعلقة ببرامج سوق العمل النشطة للشباب، مصنفة بحسب القيود التي تحاول مواجهتها. وقد توصلوا إلى أن فجوات المعرفة تكون كبيرة بشكل خاص في تدخلات مثل الفرصة الثانية في التعليم والتدريب على المهارات السلوكية والتدريب على تنظيم المشروعات وبرامج الأشغال العامة والخدمة العامة والمساعدة في البحث عن الوظائف باللجوء إلى التقنيات الحديثة وشهادات المهارات وتمويل المشاريع الصغيرة وغيرها، بما يدعو إلى التعلم والتقييم الحثيث لكل أنواع البرامج.

ملاحظة: أنظر قسم المصادر في نهاية هذا الدليل لقائمة من تقييمات الأثر التي سبق أن تمت أو في طور التنفيذ في مجال سبل عيش الشباب.

تحد ندرة الدراسات الحثيثة بشدة من وجود استثمارات واسعة النطاق في القطاع - على الرغم من الطلب الهائل على مثل هذه الدراسات. ويشكل نقص الأدلة عائقاً يحول دون كسب الدعم العام للتدخلات المتعلقة بسبل عيش الشباب. حيث يرغب المسؤولون الحكوميون عادة بتقديرات للأثر وتقديرات للتكلفة والمنفعة قبل الاستثمار في برامج كبيرة. ونتيجة لنقص هذه الأدلة في مجال سبل عيش الشباب، يكون من الصعب غالباً تقديم قضية مقنعة مقارنة بتدخلات أخرى، مثل تطوير البنية التحتية حيث تتوفر أدلة أكثر بكثير. ولهذا فإن تحسين قاعدة الأدلة من شأنه أن ييسر أيضاً توسيع نطاق التدخلات وتكرارها.

وقد أصبح هذا الاحتمال واضحاً في مجالات السياسة الأخرى. حيث مكنت الأدلة المتزايدة في برامج التحويلات النقدية المشروطة (CCT) على سبيل المثال المجتمع الدولي من تعزيز تدخلات واسعة النطاق في هذا المجال في جميع أرجاء العالم. وكانت المكسيك والبرازيل دولتان من أربعة دول فقط حول العالم يوجد فيهما برامج التحويلات النقدية المشروطة في عام ١٩٩٧، ولكن أدت الأدلة الناتجة عن تقديرات الأثر الخاصة بهما إلى توسع هائل للنموذج في ثمان وعشرين دولة في عام ٢٠٠٨ (انظر شكل ٣-١).

شكل ٣-١ من الدليل إلى السياسة: برامج التحويلات النقدية المشروطة



المصادر: فيزيين وسشادي (٢٠٠٩)؛ راولينغ وروبيو (٢٠٠٥).

سيمكن البناء المنهجي للأدلة حول الأمور التي أثبتت تأثيرها في تعزيز الفرص الاقتصادية للشباب من تحسين فاعلية وكفاءة عملنا عبر تقديم المعرفة الحيوية لخدمة الممارسين وصناع السياسات وتعزيز المجال بأكمله في نهاية المطاف.

المساءلة

بالإضافة إلى تعزيز التعليم الداخلي والتعليم في القطاع بأكمله، يدعم التقييم شرعية عملياتنا. يتزايد طلب وكالات التمويل والمجتمع على مساءلة برامج التطوير، ويمكن أن توفر عمليات التقييم - عمليات تقييم الأثر على وجه التحديد - الأدلة المطلوبة فيما إذا كان برنامج محدد قد حقق نتائج المنشودة.

مبادرات التدريب على وظيفة والتي تمولها الحكومة في أمريكا اللاتينية، وبرامج حركة الشباب المناصر Jóvenes، والتدريب التقني المتزامن مع التدريب على المهارات الحياتية وبرامج تدريب الطلاب الميداني وغيرها من الخدمات. أظهرت تقييمات الأثر آثاراً قابلة للقياس، مثل زيادة معدلات التوظيف والأجور والتي بلغت أكثر من ١٠ بالمائة بالنسبة لفئة الأصغر سناً والنساء في بعض الدول. وبالنتيجة، تم التكرار السريع لبرامج Jóvenes في جميع أنحاء القارة. وبما أنه قد تم تقييم البرامج الأحدث، أوضحت الحالة تنوع الآثار التي يمكن أن تحدث في بلدان وأوضاع مختلفة، مما يؤكد الحاجة الدائمة لقدر كبير من التقييمات بحيث يمكن تعميم النتائج. وعلى الرغم من أن تجربة Jóvenes لا تزال استثناءً في قطاع سبل عيش الشباب وذلك فيما يتعلق بالتقييم المنهجي للأثر، إلا أنها تعطي فكرة عن إمكانيات توسيع نطاق المجال إذا تمكنا فقط من تجميع معرفة وأدلة أفضل من مئات التدخلات التي نقوم بتنفيذها كل سنة.

يأخذ المانحون الدوليون بعين الاعتبار تقييمات الأثر الحثيثة بشكل متزايد من أجل قياس نجاح البرامج التي يمولونها. وفي عام ٢٠١١، عززت كل من وزارة التنمية الدولية في المملكة المتحدة (DFID) والوكالة الأمريكية للتنمية الدولية (USAID) من تركيزهما على النتائج. أنظر:

خطة عمل وزارة التنمية الدولية ٢٠١١-٢٠١٥،
القسم ٢ «زيادة فاعلية المعونة البريطانية من خلال تحسين الشفافية والقيمة مقابل المال»
<http://www.dfid.gov.uk/Documents/DFID-business-plan.pdf>

سياسة التقييم الخاصة بالوكالة الأمريكية للتنمية الدولية ٢٠١١
<http://www.usaid.gov/evaluation/USAIDEvaluationPolicy.pdf>

لا يجب اعتبار مواردنا أمراً مسلماً به. في جميع الحالات تقريباً، يتم تمويل مشروعاتنا من الأموال العامة أو الخاصة، مثل المساعدة التنموية الرسمية (وتكون أساساً من أموال دافعي الضرائب) أو التبرعات الخاصة. وفي كلتا الحالتين، فقد ائتمنا شخص ما، سواء بشكل مباشر أو غير مباشر، على استخدام هذا المال بأفضل طريقة ممكنة من أجل مساعدة الشباب على تحقيق حياة أفضل. وتعني حقيقة أننا مؤمنون على تطوير وتنفيذ مشروع سبل عيش الشباب أن هذا المال لن يستخدم لبناء طرق ريفية، أو لتحسين برنامج الوقاية من فيروس نقص المناعة المكتسبة/الإيدز، أو لشراء مواد دراسية. ونظراً لندرة الموارد، فإنه من الطبيعي استخدام عمليات تقييم من أجل تقديم بيان نزيه عن عملنا: كيف تم استخدام المال والنشاطات التي تم تمويلها والنتائج التي حققناها.

نحن مسؤولون عن ضمان الاستخدام الأمثل للأموال. تكون التدخلات التنموية معقدة بطبيعتها، وسيصعب توقع تحقيق نسبة نجاح ١٠٠ بالمائة. ومن ناحية أخرى، لا يقوم مشروع ما بزيادة رفاهية الناس تلقائياً مجرد أن له هدف سام. فمن أجل ضمان أن برنامجاً محدداً يعطي نتائج جيدة أكثر منها مؤذية وأن فوائد الاستثمار تفوق الفوائد الناتجة عن الاستخدامات البديلة للموارد، يجب دوماً على الممارسين أن يمنحوا الأولوية لتقييم فاعلية هذا التدخل (جونز وآخرون ٢٠٠٩).

المصادقية والاستدامة

تساعد التقييمات على زيادة شرعية المشروع وسمعة المنظمة المنفذة. ولا يتم عادة ذكر هذه الحجة في المواد المطبوعة، ولكن قد تكون عملياً ضمن أكثر الأسباب إلزاماً لإجراء تقييم للأثر، لأنها تفيد البرنامج والمنظمة المنفذة بشكل مباشر.

يمكن أن تحسن تقييمات الأثر من مصادقية وسمعة منظماتنا. ولأن التقييمات ذات الجودة نادرة، فإنها تحظى باهتمام خاص. وبالنسبة، فإن حقيقة موافقة منظمة أو مشروع ما على إجراء تقييم للأثر يدل بالفعل على توفر معايير جيدة في البرامج. وإذا أظهر التقييم نتائج جيدة، عندها يمكن أن تكون الفائدة التي تعود على المنظمة والبرنامج هائلة. تخيلوا أنه من بين مئات اللاعبين في المجال، كنتم أنتم وحدكم قادرون على إثبات أن طريقتكم ناجحة، وأن برنامجكم يوفر بنجاح فرص دخل للشباب والشابات تفوق بوضوح فرص الدخل التي قد تتوفر لهم في حالة عدم مشاركتهم في برنامجكم. يتمثل الفرق في أنه بإمكانكم الآن ليس فقط ادعاء أن تدخلكم فعال، ولكنكم قادرين على إثبات ذلك. وهذا يشكل فرقاً كبيراً من وجهة نظر المانحين وصناع السياسات، والذين لم يكونوا قادرين - قبل التقييم - على التفريق بين آثار تدخلكم وبين الآثار المزعومة للبرامج العديدة الأخرى.

يمكن للقدرة على التمييز أن تمنح سلسلة من الفوائد للمشروع والمنظمة على حد سواء. حيث يمكن استخدام نتائج التقييم الإيجابية في جهود حشد التأييد والتمويل من أجل الحصول على دعم أكبر من المانحين والحكومات وعامة الشعب. ومع دعم شعبي وسياسي أكبر، يمكن لمشروعنا ومنظمتنا أن يصبحنا بسرعة مرجعاً في المجال. وهذا بالمقابل يؤدي غالباً إلى زيادة في الطلب على الخدمات، وقد يكون من المتوقع منا توسيع خدماتنا على الصعيد الوطني وعبر الحدود.

لنأخذ على سبيل المثال حالة كوليكتيفو انتيغال دي ديسارولو Colectivo Integral de Desarrollo في بيرو. ففي عام ٢٠٠٣، كانت المنظمة ضمن أوائل المنظمات في تقديم أدلة حثيثة على أن نموذجها القائم على تعزيز رواد الأعمال الشباب ذوي الدخل المنخفض كان يؤدي إلى زيادة حجم العمل التجاري وتحسين فرص استمراريته وزيادة الدخل. ونتيجة لهذا الدليل، حصلت المنظمة على عدة جوائز ولم تعد تعاني من مشاكل في تأمين التمويل لبرامجها. ففي الواقع أن كوليكتيفو مدعومة الآن ببرنامج منحة من مصرف التنمية للبلدان الأمريكية وتعمل على توسيع نموذجها إلى أمريكا الوسطى والكاريبي (انظر مربع ١-٢).

أتى جواب رئيس المنظمة غير الحكومية في البيرو كوليكثيفو انتيغفال دي ديسارولو Colectivo Integral de Desarrollo على سؤال «برأيك، كيف استفادت منظمتك من إجراء تقييم للأثر؟»

«حسن من نوعية تدخلنا»

«حسن من مصداقية البرنامج»

«حسن من قيمة علامتنا التجارية من وجهة نظر المانحين»

«زاد من الطلب على خدماتنا»

«اكتسبنا الاعتراف المحلي والدولي»

«نحن مؤسسة نموذجية يمكن تكرارها في سياقات وبلدان أخرى»

المصدر: دينو ليناريس، رئيس كوليكثيفو انتيغفال دي ديسارولو Colectivo Integral de Desarrollo، اتصال شخصي (٢٨ يناير/ كانون الثاني، ٢٠١١).

نتائج التقييم السليبي ليست بالضرورة سيئة

ربما نخشى من أن تؤدي النتائج السلبية من التقييم إلى قطع التمويل من مانحيننا، إلا أنه يمكن الاستفادة من التقييمات التي تفشل في إثبات نتائج إيجابية لتدخلاتنا.

لا يمكن تجنب النتائج السلبية في البرامج المشتركة. إن الابتكار والإبداع أمران حاسمان في مساعدة الشباب على إتقان انتقالهم نحو بيئة العمل. وسينطوي هذا الابتكار بالتأكيد على حالات من الفشل. وكما الحال في أي مجال آخر مثل الطب أو الكيمياء أو الفيزياء، يتطلب بناء منتجات وخدمات ناجحة الاختبار والنماذج التجريبية والصقل والتكيف مع الظروف المحلية. وتمثل حالات الفشل خطوة ضرورية باتجاه وضع أحدث البرامج.

يمكن أن تساعد النتائج السلبية في تحسين العمليات. إذا تمكنا في وقت مبكر من فهم المشاكل التي يمكن أن تقلل من فاعلية تدخلاتنا، نكون عندئذ في وضع جيد من أجل بناء مشروعات ناجحة على المدى الطويل. حيث توجهنا الأخبار السيئة من نتائج التقييم السلبية نحو سبل تحسين وضع برامجنا.

تعزز المعالجة الاستباقية للنتائج السلبية المصداقية. لا يتوقع أو يصدق مانح أو صانع سياسة أن كل مشروع سيحقق نجاحاً كبيراً. إن نشر النتائج، سواء الإيجابية أم لا، يشير إلى قدرتنا على النقد الذاتي والتزامنا بالتعلم المستمر وبناء البرامج التي تعتمد على الأدلة. ومن المسلم به أن الضغط الذي تتعرض له البرامج لإظهار نتائج لتدخلاتها وتبرير الميزانيات التي تم انفاقها يمكن أن ينشئ دوافع قوية للإبلاغ عن نتائج إيجابية أكثر من النتائج السلبية. ولكنه على المدى الطويل، يحتمل أن يثمر النقاش الصادق حول الأمور التي نجحت والأمور التي لم تنجح عن الفائدة الأكبر.

النقاط الرئيسية

١. تتمحور التقييمات في البداية وقبل كل شيء حول التعلم من أجل فائدة مشروعنا ومنظمتنا. حيث تسمح لنا التقييمات بإظهار القيمة الحقيقية لعملائنا مما يعود بالفائدة على تصميم وتخطيط التدخلات الأخرى.
٢. تعطي التقييمات قاعدة أدلة وهي حاجة ملحة في مجال سبل عيش الشباب. وعليه فإن المعرفة الأكثر والأفضل عما يجدي نفعاً وما لا يجدي ستخدم الممارسين في تصميم تدخلات ناجحة وإقناع صناع السياسات بمنح دعم عام.

«لأننا تمكنا من برهنة قدرتنا على تعزيز فرص توظيف الشباب، قامت الحكومة بدعوتنا للمشاركة في تصميم برنامج لتعزيز فرص التوظيف تحت إشراف الرئيس الجديد.»
— مدير البرنامج، منظمة تشيلية غير حكومية.

قدم برنامج في الشرق الأوسط نهجاً مبتكراً في تدريب شباب على وظائف مساعدات تنفيذيات لصاحبات الأعمال. وقد أمد البرنامج قطاع الأعمال التجارية وصاحبات الأعمال على وجه التحديد بصلات مهمة، حيث كن مهمات بدعم وتمكين الشباب الخرومات. ولكن وجد التقييم الميداني مستويات عالية نسبياً من التسرب للتدريبات الشباب بعد أن تم إيجاد عمل لهم. ووجدت المزيد من التحقيقات حول سبب هذا التسرب أن صاحبات الأعمال قد عقدن أمالاً كبيرة على هؤلاء الشباب، ولكنهن لم يوفرن لهم الإرشاد الكافي من أجل دعم نجاحهن في بيئات العمل المتطلبية. وعبر الحوار الصريح مع الشريك المحلي المنفذ، أعاد البرنامج صياغة نماذج المهارات الحياتية من أجل تحضير الشباب بشكل أفضل لمتطلبات هذه الوظائف، كما عمل على إعادة توجيه أصحاب الأعمال من أجل ضمان تقديمهم الإرشاد الكافي للموظفات الجدد. وتمكنت الوكالة المنفذة من تحقيق نتائج أفضل بسبب هذه التعديلات أثناء العمل والتي تم تطبيقها على استراتيجية البرنامج، وتمت مشاركة هذه الدروس مع المانحين، وبالمقابل، أمنت تمويلًا إضافيًا من أجل توسيع نطاق البرنامج.

- قوة قياس النتائج
- إذا لم تقوموا بقياس النتائج، فلن تتمكنوا من تمييز النجاح من الفشل.
 - إذا لم تتمكنوا من رؤية النجاح، فلن يمكنكم التعلم منه.
 - إذا لم تتمكنوا من رؤية الفشل، فلن تتمكنوا من تصحيحه.
 - إذا كان بإمكانكم إثبات النتائج، يمكن أن تفوزوا بدعم وتمويل عام.
- المصدر: مقتبس عن أوزبورن وغيلبر (١٩٩٢).

٣. توفر التقييمات الشرعية وذلك عبر اعتبار أنفسنا مسؤولين أمام المانحين والجمهور. حيث تضمن التقييمات الاستخدام الجيد لأموال دافعي الضرائب والتبرعات.
٤. تحسن التقييمات من مصداقيتنا وسمعتنا. ففي مجال تندر فيه الأدلة الحثيثة، يمكن أن يكون لإجراء التقييمات فوائد كبيرة فيما يتعلق بزيادة الطلب على خدماتنا، وتعزيز العلامة التجارية لمنظمتنا، وتأمين دعم مالي مستدام في نهاية المطاف.
٥. لا يتعين أن تظهر التقييمات نتائج جيدة لكي تكون مفيدة. فعلى العكس، يعزز الفشل التعلم. وستؤدي المعالجة الاستباقية لنتائج التقييم السلبية ونشرها على الأغلب إلى تحسين مصداقيتنا وسمعتنا.

دراسة حالة صندوق العمل الاجتماعي لشمال أوغندا (NUSAF): لماذا التقييم؟

قامت الحكومة الأوغندية بإطلاق تقييم لـ NUSAF بدعم من البنك الدولي. وكان السبب الرئيسي لتقييم الأثر هو تحسين إدارة البرنامج. وسعيًا لتقييم أثر برنامج فرص الشباب على سبل عيش ورفاه الشباب في شمال أوغندا، كانت الغاية من التقييم توفير معلومات لبرمجة المراحل المستقبلية وإمكانية توسع حجم المشروع.

كما كانت الغاية من تقييم الأثر هي ملء فجوة مهمة في فهم فاعلية برامج التوظيف والتدريب على مهارات تنظيم الأعمال، في الإطار الإفريقي على وجه الخصوص. وعبر تقديم منح من أجل الحصول على تدريب على المهارات ورأس مال ابتدائي لتأسيس مشروعات منتجة، فإن برنامج فرص الشباب هو مزيج من نوعين من أكثر الأنواع شيوعاً في برامج التوظيف. وبما أنه لا يعرف سوى القليل عن فاعلية مثل هذا النهج، فإن التقييم سيولد المعرفة التي يمكن أن تثري مجال سبل عيش الشباب بالكامل.

المصدر: باتلمان ومارتينيز وفيالا (٢٠٠٩).

قراءات رئيسية

سيفدوف، ديليو وليفين آر ويردسال إن. ٢٠٠٦.

When Will We Ever Learn? Improving Lives through Impact Evaluation

واشنطن دي سي: مركز التنمية العالمية.

<http://www.cgdev.org/content/publications/detail/7973>

بنود

بند ٢: مراجعة تصميم المشروع

الشخص الوحيد الذي اعرفه ويتصرف به عقل
هو الخياط فهو يأخذ مقاساتي من جديد في كل مره يراني ...
أما الباكون فيستخدمون مقاييسهم القديمة ويتوقعون
مني ان اناسبها!!

— جورج برنارد شو

تخيل نفسك في المواقف التالية:

طلب محافظ طرابلس بلبنان من مؤسستك تقديم الدعم الفني من أجل معالجة ظاهرة البطالة المنتشرة بين شباب المدينة. ما هو التصرف السليم الذي يجب القيام به؟

أبدى مدير مكتب منظمته في رواندا اهتمامه ببرامج عمل الشباب لحسابهم الخاص. في رأيك، ما هي الإجراءات التي تقترح اتخاذها؟

خطتلك الاستراتيجية للأعوام الثلاث المقبلة تركز بشكل أكبر على توجه الشباب نحو العمل. ما هو مخطط توظيف الشباب الذي ينبغي عليك الاستثمار فيه؟

يناط بمدراء البرامج اتخاذ قرارات صعبة بشأن هذه القضايا وغيرها من القضايا المتعلقة بوضع البرامج. لذا كي يتسنى لنا إيجاد الحلول المناسبة، نحتاج إلى فهم السياق المحدد وتصميم برنامج يتفق مع هذا السياق. ومن ثم، قبل الشروع في صياغة نظام المتابعة والتقييم، نحتاج إلى التأكد من أن التدخلات تم تخطيطها بشكل دقيق: هل لدينا المعرفة الجيدة حول احتياجات الأشخاص الذين نسعى إلى دعمهم؟ هل ندرك سبب انتشار هذه الظواهر كبطالة الشباب والعزل الاجتماعي؟ هل لدينا هدف واضح بشأن ما نرغب في تحقيقه؟ وأخيراً هل نركز على خبرات ودلائل قائمة بالفعل عند تصميم تدخلنا لتحقيق هذا الهدف؟

تحليل المشكلة: هل نفهم المجموعة المستهدفة والسياق المحلي؟

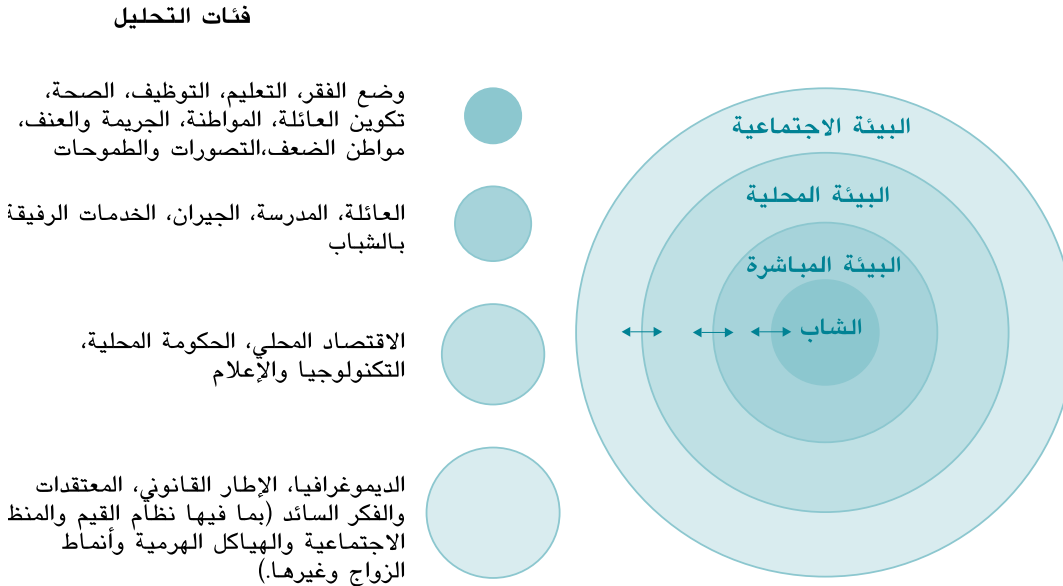
عندما نقرر التدخل من أجل دعم الشباب، نفترض ضمناً وجود تناقض بين الوضع الراهن وما كان يمكن أو يجب أن يكون. هذه الفجوة بين الوضع القائم والوضع المرغوب هو ما نشير إليه بالاحتياج. لذا يتعين علينا فهم هذه الاحتياجات قبل البدء في تطبيق برنامج التدخل.

استخدام تقيييمات الشباب متعددة القطاعات من أجل فهم مجموعتنا المستهدفة

يصعب فهم الفرص الاقتصادية إذا ما تم تقييمها بشكل معزول، حيث إنها مرتبطة ارتباطاً وثيق الصلة بمناحي الحياة المختلفة. فعلى سبيل المثال، قد يحدد الوضع الوظيفي والدخل قدرة الشخص على الزواج أو تكوين عائلة. لذا فإن الفهم الواعي لمشاكل الشباب واحتياجاتهم يحتاج على الدوام إلى إجراء تقييمات متعمقة للظروف المعيشية للشباب، والتي تشمل على وضعهم الاجتماعي والاقتصادي وعلى سلوكياتهم ومواقفهم فضلاً عن غاياتهم وتطلعاتهم. كما ينبغي أن تتضمن تقييمات الشباب بعض العوامل الهامة الأخرى ذات الصلة بانتقال الشباب إلى سن الرشد، مثل الأمور الصحية وتكوين العائلة وواجبات المواطنة (لمزيد من المعلومات حول الانتقالات إلى سن الرشد، يرجى زيارة رابط البنك الدولي ٢٠٠٦).

وعلاوة على ذلك، حيث يتأثر الشباب بالعديد من العوامل المحيطة بهم، بما في ذلك العائلة والأقران والمجتمعات والمؤسسات المحلية والوطنية فضلاً عن المعايير الاجتماعية، يتعين على تقييمات الشباب الجيدة أن تتجه كذلك إلى تحليل البيئة المباشرة والمحلية والاجتماعية التي يعيش فيها الشباب (أنظر برونفينبرينر ١٩٧٩). إن التقييم الشمولي سيرسم صورة مفصلة للاحتياجات والتحديات التي يواجهها الشباب والتي سوف تسمح لنا بدورها بتكييف عملية تدخلنا بشكل أفضل مع الوقائع على المستوى المحلي (أنظر شكل ٢-١). تم إدراج قائمة تحتوي على عينة من المؤشرات الخاصة بتقييم الشباب في ملحق ١.

شكل ٢-١ بيئات الشباب



عند إجراء تقييم سريع للمجتمع عن الخصائص الاجتماعية الاقتصادية للشباب المستهدف في الأردن، لاحظت المؤسسة الدولية للشباب الانخفاض الشديد في مستوى مشاركة الشباب في الأنشطة المدنية في اثني عشر من المجتمعات المستهدفة (أقل من ٤٪). علاوة على ذلك، كشفت بيانات المسح عن وجود قلة قليلة من المؤسسات التي تقدم فرص تطوعية. فعندما يشارك الشباب في مشاريع خدمات مجتمعية، يكون ذلك في المقام الأول من خلال مدارسهم. وضحت حلقات النقاش البؤرية المركزة التي أجريت مع مجموعة من الشباب أنه على الرغم من أن «ثقافة العمل التطوعي» لم تطرح في هذه المجتمعات، أبدى الشباب حماسهم واستعدادهم للتطوع حال توفر الفرص لذلك.

وقد خدمت هذه النتائج تصميم مشروعات تعلم خدمية محددة للعاطلين عن العمل الذين تسربوا من المدارس. كما أنها بررت تقديم المنح للشباب للقيام بمبادرات مجتمعية صغيرة، الأمر الذي جعل خيارات المشاركة المدنية ظاهرة ومتاحة بشكل أكبر وضمنت أنها تجذب الشباب. بالإضافة إلى ذلك، دربت المؤسسة الدولية للشباب العاملين في منظمات الشباب الخدمية على ماهية التنمية الفعالة فضلاً عن إدارة البرامج المعنية بالمشاركة المجتمعية والبرامج التطوعية، والتي ساعدتهم في إشراك الشباب بشكل أفضل داخل مجتمعاتهم.

سيساعدنا تحليل البيانات الشخصية والاجتماعية للشباب من خلال إجراء مقابلات منتظمة وتشكيل مجموعات بؤرية/تركيز والملاحظة في تحديد المشكلة الرئيسية التي نسعى إلى معالجتها، مثل نقص التوظيف أو البطالة أو العجز عن الوصول إلى الخدمات المالية. وقد ندرك كذلك أن الفرص الاقتصادية المحدودة إنما هي واحدة من بين العديد من القضايا التي يواجهها الشباب في أماكن محددة، الأمر الذي قد يوحي بدوره بسبل لإرساء أو تكييف عملية التدخل من طرفنا بحيث يمكنها معالجة أكثر من قضية واحدة.

وعلى نفس القدر من الأهمية، سوف يساعدنا التقييم على تحديد مجموعتنا المستهدفة. فهل نحن مهتمون بالشباب كافة، أم هؤلاء الذين انقطعوا عن التعليم؟ ما هي الفئة العمرية التي نرغب في التركيز عليها؟ هل ثمة اعتبارات للجنس أو العرق نود أن نجعلها ذات أولوية؟ ما هي المناطق الجغرافية التي نستهدفها؟ نظراً للقيود على مواردنا، يصعب علينا خدمة كل شاب. لذا فإن تقييمات الشباب المتعددة القطاعات يمكن أن تساعدنا في تحديد الأولويات، على سبيل المثال من خلال تحديد المجموعات الضعيفة بشكل خاص، مثل حالات التسرب الدراسي أو الفتيات الشابات أو شباب الشوارع.

استخدام تقييم السوق لفهم الاقتصاد المحلي

نظراً لتركيزنا على بناء أو تعزيز سبل العيش، يعد تقييم الاقتصاد المحلي أحد الركائز الأساسية في السياق التحليلي. حيث يشمل تحليل الاقتصاد المحلي عادة على تقييم سوق العمل المحلي وتقييم سوق السلع والخدمات.

تقييم سوق العمل المحلي

تنشد عمليات تقييم سوق العمل فهم أنماط واتجاهات عملية التوظيف في الاقتصاد المحلي. وتشمل العوامل المشتركة التي تخضع للتحليل خلال إجراء ذلك التقييم ما يلي (بنك التنمية الآسيوي ٢٠٠٧، ص. ١٦٢-١٦٦):

- **الطلب على العمالة:** الظروف الاقتصادية الشاملة؛ حجم القطاعات الرسمية وغير الرسمية؛ القطاعات أو الصناعات الديناميكية والمناطق الجغرافية التي تطلب عمالة؛ اتجاهات الصناعة وتوقعاتها؛ فرص العمل المتوقع خلقها؛ المهارات المطلوبة حسب فرص العمل الشاغرة؛ مستوى الأجور والدخول؛ ظروف العمل؛ ممارسات التعيين؛ تصورات أرباب العمل؛ عوائق التوظيف على أساس الجنس أو العمر (السن) أو العرق أو الوضع الاجتماعي أو الديانة أو غيرها من الأسباب الأخرى؛ وما إلى ذلك.
- **عرض العمالة:** حجم وخصائص السكان في سن العمل؛ التوظيف ونقص فرص العمل والبطالة حسب الجنس والعمر (السن) والمستوى التعليمي والمناطق الحضرية/الريفية والقطاع الاقتصادي وفرص العمل الشاغرة والقطاعات الرسمية/غير الرسمية والقطاع العام/الخاص.
- **البيئة المؤسسية وبيئة السياسات.** برامج سوق العمل القائمة والسياسات والقوانين والمؤسسات بما في ذلك، على السبيل المثال، الحد الأدنى للأجور وقوانين حماية التوظيف وتنظيم الاتحادات النقابية وإعانات البطالة وغيرها. ومن بين الجوانب الأخرى المثيرة للاهتمام أولويات القطاعات الاقتصادية والتي يتم تحديدها على المستويات القومية والإقليمية والمحلية.

تقييم سوق السلع والخدمات

إن تقييم سوق السلع والخدمات من شأنه أن يساعد في تحديد إمكانية مشاركة صغار المنتجين في الأنشطة الاقتصادية المستدامة والتوزيع المحتمل للأدوار (على سبيل المثال، دور الشباب أو المرأة) في هذه الأسواق (بيروز-بكلي ٢٠٠٧). وتشمل خصائص السوق المشتركة المقرر تحليلها:

- **طلبات السوق وسلاسل القيمة.** الفجوات الحالية والمستقبلية فيما يتعلق بالمنتجات والخدمات الاستهلاكية؛ الطلب على السلع والمنتجات المعالجة والسلع الوسيطة من قبل تجار التجزئة والجملة أو الشركات المصنعة، تحديد الأسواق المحلية والإقليمية وأسواق التصدير؛ تحديد رواد السوق الحاليين بالإضافة إلى غيرها من العوامل الأخرى.
- **استقرار السوق.** مواطن ضعف السوق نحو الصدمات واتجاهات التغيير والتغيرات الموسمية والقيود المحتملة للوصول إلى السوق وحركة الأفراد والمنتجات نتيجة الصراعات وعدم الأمن.

- **أسعار السوق.** تقلب أسعار المنتج النهائي والمواد؛ الأثر المحتمل لدخول منتجين إضافيين على الأسعار، والتضخم وتكاليف المعاملات التجارية.

تتم تقييمات السوق عادة من خلال الجمع بين تحليل البيانات المتاحة وإجراء مسح لأرباب العمل أو أصحاب المشروعات الصغيرة. يمكن للأفراد الذين تتم مقابلتهم تقديم رؤى هامة حول احتمالات التوظيف في قطاعات محددة وكيفية اتخاذ قرارات التعيين والقيود الرسمية وغير الرسمية الهامة التي تواجه الأعمال التجارية وتصوراتهم عن الشباب وما إلى ذلك. يمكن لهذه المعلومات تباعاً أن تشخص وتوفر معلومات لتصميم البرنامج (أنظر الأقسام التالية). أنظر مربع ١-٢ حول أمثلة التقييمات التي أجريت على الشباب والسوق.

مربع ١-٢ أمثلة التقييمات التي أجريت على الشباب والسوق

التقييم المتعمق للشباب والسوق:

ببترز وآخرون ٢٠٠٩. *Youth Employment in Sierra Leone*. واشنطن، دي سي: البنك الدولي.
http://www-wds.worldbank.org/external/default/WDSPContentServer/WDSP/IB/2009/02/27/000334955_20090227091204/Rendered/PDF/476090PUB0Sier101OfficialUse0Only1.pdf

مسح سريع لتقييمات الشباب:

المؤسسة الدولية للشباب. ٢٠١٠. *Building on Hope: Findings from a Rapid Community Appraisal*. in Jordan. بالتييمور: IYF.
<http://www.iyfnet.org/document/1059>

المؤسسة الدولية للشباب. ٢٠١١. *YouthMap Senegal— Youth Assessment: The Road Ahead*. بالتييمور: IYF.
<http://www.iyfnet.org/document/1820>

تركز مبادرة LEGACY الخاصة باللجنة الدولية للإنفاذ (IRC) في ليبيريا على دعم البرامج التعليمية الموجهة مجتمعيًا وتوسيع نطاق فرص التدريب المهني التي يوجهها السوق للفتيات والشباب المعزولين. ولضمان أن المنهج المتبع كان موجهًا بالسوق، أجرت لجنة الإنفاذ الدولية تقييمًا.

حددت دراسة استقصائية لسوق العمل المهن الرائجة والتي تتوافر فيها فرص عمل محتملة للشباب. وتمثلت أداة التقييم في مسح مكون من عشر صفحات يحتوي على أسئلة اختيار من متعدد وأسئلة مفتوحة تهدف إلى فهم أفضل لاحتياجات التوظيف القائمة والمحتملة لأرباب العمل.

بالإضافة إلى ذلك، أجرت IRC دراسة استقصائية سريعة للسوق لمساعدة اثنين من مراكز التدريب المهني في تحديد المنتجات والخدمات للأعمال التجارية المحتملة المستندة إلى المدارس. اشتمل التقييم إجراء مسح مع تجار التجزئة والمستهلكين، والموردين في الأسواق المحلية.

ساعدت التقييمات IRC في فهم احتياجات الفتيات مقارنة باحتياجات الشباب، الأمر الذي ساعد في النهاية بدوره في ارتفاع عدد الفتيات المتحقات بمراكز التدريب.

المصدر: بيوفي-ساني وآخرون (٢٠٠٩).

التشخيص: ماهي المحددات المؤثرة على نتائج الشباب؟

ماذا بعد التقييمات التي تجرى على الشباب والسوق؟ فلنفرض أن الشباب في البلد أو المنطقة أو المدينة التي نعمل فيها يتأثرون بشكل متفاوت بالبطالة ونقص فرص التوظيف. أكثر الشباب ممن تتراوح أعمارهم بين ١٦-٢٤ عاماً غير ملتحقين بالتعليم ولا يعمل. ويبدو أن الفتيات تحديداً أكثر تأثراً. ما هي خطواتنا التالية. وما هو التدخل الذي يمكن أن نقترحه؟ في الواقع، تعتبر هذه الأسئلة سابقة لأوانها. فقبل أن نفكر في التدخل، نحتاج إلى أن نعرف لماذا يعجز هؤلاء الشباب عن إيجاد عمل أو بدء عمل تجاري خاص بهم. ما السبب الذي يمنعهم من الدخول إلى سوق العمل وتوفير سبل العيش لهم ولعائلاتهم؟ ما هي القيود التي يواجهونها؟ تحليل أننا وضعنا دورات تدريبية في المجالات الفنية تستهدف الشبابات. قد تقدم الدورات التدريبية تعليم للمهارات الفنية والحاسب الآلي وذلك على أساس نموذج قامت منظمنا بتنفيذه بنجاح في مكان آخر. لكن ماذا إذا لم تكن المهارات الفنية هي المشكلة التي يجب البدء بها؟ ماذا إذا كانت المشكلة الحقيقية لهؤلاء الشبابات الافتقار إلى المعرفة بشأن كيفية البحث عن عمل والتقدم له، فضلاً عن القيود الاجتماعية التي تثبط من عزيمتهن الشبابات من العمل خارج نطاق المنزل؟ فإذا كان لدينا قدر ضئيل من المعرفة أو وضعنا الافتراضات الخاطئة بشأن السبب الكامن وراء عدم قدرة الشباب على إيجاد فرص عمل، فلن يعالج مشروعا المشكلة الأصلية ومن ثم لن يكتب له النجاح. وفي هذه الحالة، سوف تؤكد عملية المتابعة والتقييم الحقائق الواضحة.

فهم الحواجز طويلة المدى والحواجز قصيرة المدى

يتحتم فهم القيود الضمنية التي قد تحد من دخول الشباب سوق العمل وفي الأنشطة التي تحقق دخلاً. ونحن هنا نحاول وضع مفاهيم عن هذه القيود بالإضافة إلى القيود التي يحتمل عدم إمكانية تناولها خلال الإطارات الزمنية لأغلب البرامج (طويلة المدى) وكذلك التي يمكن تناولها (قصيرة المدى). **القيود طويلة المدى.** تستغرق القضايا المؤسسية وقضايا الاقتصاد الكلي وقتاً في تناولها ولا يرجح تأثرها بالمشروعات المحلية الفردية التي يتراوح الإطار الزمني لها في المعتاد بين ثلاث وخمس سنوات إلا أنه من المهم دراسة القيود طويلة المدى لأنها تمثل السياق الأكبر للتدخل الخاص بنا. يعرض مربع ٢-٢ إطار عمل مايلز (MILES)، نظرة عامة عن المحددات الهيكلية خلق الوظائف.

مربع ٢-٢ إطار عمل مايلز (MILES)

الاقتصاد الكلي والاستقرار السياسي. يحتاج أصحاب الأعمال إلى إطار اقتصادي كلي سليم للتوسع في أعمالهم وإيجاد فرص عمل جديدة.

مناخ الاستثمار والمؤسسات والبنية التحتية. ستتوسع الشركات وتوجد وظائف في القطاعات الرسمية عندما تكون تكاليف إنشاء الأعمال (من اللوائح وأعباء الضرائب الثقيلة والبنية التحتية الضعيفة) منخفضة ويمكن التنبؤ بها.

لوائح ومؤسسات سوق العمل. اللوائح السليمة حتمية لكل من رب العمل والعامل لإقامة علاقة عمل منتجة وطويلة المدى.

التعليم والمهارات. تعتمد الوظائف ذات الإنتاجية المرتفعة بشكل ثابت على التعليم الرسمي الجيد وتتطلب مهارات مناسبة لكافة الفئات العمرية.

الحماية الاجتماعية. إن وجود نظام حماية اجتماعية قوي ومتوازن يحمي دخل العاملين من الصدمات عند التوظيف.

(المصدر: البنك الدولي (مقتبس من ٢٠٠٧، ص. ٨-١٠)

القيود قصيرة المدى. نظراً لصعوبة تغيير غالبية الحواجز الهيكلية التي تواجه التوظيف وإيجاد سبل العيش، عادة ما يكون التركيز على القيود التي يمكن تناولها في فترة زمنية أقصر أكثر واقعية بالنسبة للعاملين في مجال التنمية. ومن بينها خمسة فئات رئيسية بارزة (كننغهام وسانشيز-بيروتو وويرملي ٢٠١٠):

- **القيود من جانب العرض:** يفتقر الشباب إلى المهارات الوظيفية المتعلقة بالسوق المحلية بما في ذلك مهارات معرفة القراءة والكتابة الرئيسية وأساسيات الرياضيات أو المهارات الفنية أو المهارات السلوكية أو مهارات إقامة مشروعات صغيرة، وقد يواجهون كذلك قيود لا تتعلق بالمهارة، بما في ذلك قضايا نفسية قد تؤثر على فرص توظيفهم (روسياسكو وآخرون ٢٠١٠).
- **القيود من جانب الطلب:** لا يقدم أرباب العمل سوى العروض القليلة للعمالة الشابة بسبب الآثار ذات المستوى الكلي مثل بطء نمو الوظائف بالإضافة إلى الآثار ذات المستوى الجزئي مثل التمييز لدى أرباب العمل.
- **قيود إنشاء العمل:** تتضمن القيود على الشباب المنظم للمشروعات عدم إمكانية حصولهم على رأس المال أو الأرض أو تشييك.
- **قيود وساطة سوق العمل:** كثيراً ما يفتقر الشباب إلى المعلومات الدقيقة وذات الصلة عن فرص الوظائف الشاغرة والمؤهلات المطلوبة أو لا يستطيعون إيصال مهاراتهم بشكل ملائم لأرباب العمل المحتملين.
- **القيود الاجتماعية:** قد تحد الأعراف أو العادات الاجتماعية من تنمية المهارات أو دخول سوق العمل بالنسبة لمجموعات معينة مثل الفتيات والشباب أولاد البلد وغيرهم.

أثناء تصميم مشروع معني بتوظيف الشباب وبرنامج تطوير مشروعات في إندونيسيا، لم تجري منظمة محلية غير هادفة للربح تشخيصاً كاملاً للمشكلات النفسية الكامنة والقيود الاقتصادية التي واجهها الشباب إثر وقوع زلزال مدمر في المنطقة في عهد سابق. ولم يسفر الزلزال عن موت الكثيرين فقط بل تسبب في تشريد أكثر من ٥٠,٠٠٠ نسمة كما أدى إلى تدمير منشآت سبل العيش. ولم تتوافر خدمات استشارية حيوية للشباب بعد تعرضهم لهذه الصدمة؛ وفي واقع الأمر، لم يتم إدراك الحاجة إلى دعم نفسي مستمر. ورغم أن الزيارات المتابعة قد كشفت عن نجاح بعض محاولات خلق فرص عمل، أخفق البرنامج في تحقيق الأهداف الشاملة التي ينشدها حيث أن الشباب استمروا في معاناتهم من الاكتئاب والصراع لتبني الأساليب التقنية الجديدة أو أخذ خطوات أخرى قد تجعل مشروعاتهم الصغيرة أكثر ربحية.

يقدم جدول ٢-١ نظرة عامة على القيود الممكنة. بشكل عملي يكمن التحدي في تحديد أيها أكثر صلة بسياقنا المحلي وبالتالي منحها الأولوية. كل قطاع فرعي من السكان محل الاهتمام سيواجه على الأرجح مجموعة مختلفة من القيود. على سبيل المثال المرأة الشابة التي تعيش في ريف رواندا في اقتصاد منخفض النمو ستفتقر إلى المهارات وستواجه تمييزاً هائلاً في المعاملة من رب العمل وستحدها أعراف النوع في حين أن الرجل منخفض الدخل في المناطق الحضرية في تشيلي قد يكون أكثر تقيداً بالمعلومات عن فرص العمل ويواجه صعوبة في توصيل كفاءاته لأرباب العمل المحتملين وعدم توافق المهارات الفنية أو الإنسانية (كننغهام وسانشير-بيرتو وويرملي ٢٠١٠). ويتعين تحديد القائمة القصيرة للقيود الخاصة بالسكان المستهدفين المحددين من خلال البيانات التحليلية للشباب وللسوق كما هو مبين أعلاه.

جدول ١-٢ نظرة عامة على القيود قصيرة المدى التي تواجه الشباب في سوق العمل

القيود	الوصف	مصادر البيانات
جانب العرض: المهارات المتعلقة بالوظيفة وحوافز أخرى من جانب العرض	• معرفة القراءة والكتابة وأساسيات الرياضيات هما أساس التواصل وكذلك عمليات تنمية مهارات أخرى. • الشباب في الأماكن التي مرت ب نزاعات قد يتأثرون بشكل خاص.	• تقييمات المهارات متضمنة شهادة المهارات (مثل البرنامج الدولي لتقييم الطلبة PISA والاتجاهات العالمية في دراسة الرياضيات والطوم TIMSS). • تقييمات النظام التعليمي (SABER) • إحصائيات التحصيل التعليمي • المنهج المدرسي (ما الذي تم تعليمه؟) • مسح أرباب العمل الحاليين • تقييمات السوق/القطاع
	• تتراوح المهارات البيعية أو الوظيفية بين مهارات يدوية إلى معرفة استخدام الكمبيوتر.	
	• المهارات السلوكية – أو المهارات الحياتية – تتكون من تشكيلة من الخصائص مثل امتلاك الدافع وحل المشكلات والاتصال والتواصل وإدارة الوقت والقدرة على العمل مع الآخرين. • تزداد قيمة المهارات السلوكية لدى أرباب العمل حول العالم.	
	• القدرة الإبداعية لاختراع أو تبني منتج جديد أو عملية جديدة ومهارات العمل لتسويق الفكرة عوامل أساسية بالنسبة لكل من الموظفين ومن يعملون لحسابهم الخاص.	
	• قيود أخرى قد تؤثر على قدرة الشباب على قبول العمل أو البحث عنه مثل تكلفة وسائل المواصلات ومسؤولية رعاية الطفل وغيرها. • يمكن أن تؤثر صحة الشاب، خاصة الصحة العقلية، على فرص توظيفه. على سبيل المثال، يمكن للاكتئاب أو القلق أن يؤثرًا على مهارات السلوكية.	• تقييم الشباب • تقييمات أو دراسات متخصصة للصحة والصحة العقلية

جدول ٢-١ نظرة عامة على القيود قصيرة المدى التي تواجه الشباب في سوق العمل

المقيّد		الوصف	مصادر البيانات
جانب الطلب: عدم وجود طلب على العمالة	بطء نمو الوظائف	• قلة الوظائف الجديدة للغاية (قطاع رسمي صغير). • النمو البطئ ككبداً ما ينتج عن عوامل على المستوى الاقتصادي مثل بيئة الاستثمار الصعبة أو من الصدمات الخارجية مثل الكوارث الطبيعية أو الحروب أو التغيير المفاجئ في الاقتصاد العالمي.	• بيانات الناتج المحلي الإجمالي (حسب القطاع) • إحصائيات سوق العمل (بما في ذلك حصة من الاقتصاد الرسمي/ غير الرسمي) • مسح أرباب العمل الحاليين • تقييم السوق
	تميز في المعاملة من رب العمل	• قد يكون لدى أرباب العمل تعامل ضد الشباب كأن يكونوا مقتنعين على سبيل المثال بأن الشباب أقل فاعلية وأقل جدارة بالثقة أو أقل كفاءة من الأشخاص الأكبر سناً. • بالإضافة إلى التعامل قد يفضل التعيين على أساس نوعي أو عرقي أو ديني.	• مسح أرباب العمل الحاليين • تقييم السوق • دراسات علم الاجتماع
	عدم سهولة الوصول إلى رأس مال مادي وطبيعي واجتماعي	• قيود على العمل للحساب الخاص قد تتضمن الاقتراض إلى مهارات إقامة الأعمال بالإضافة إلى الحصول على الأموال أو الأرض أو شبكات الأعمال بطريقة غير ملائمة.	• تقييم السوق • الإحصائيات البنكية/إحصائيات التمويل الأصغر (منتجات القروض، متطلبات الضمانات وغيرها) • حقوق الملكية
	تأسييس الأعمال: قيود لقيام الشركة	• غالباً ما يقتصر الشباب إلى التشبيك لإيجاد الوظائف المتاحة.	• تقييم الشباب • تقييم السوق
القيود الاجتماعية	انكفاءات البارزة	• قد يمتلك الشباب المهارات الصحيحة ولكن قد يكون من الصعب إيصالها هذه المهارات إلى أرباب العمل المحتملين (على سبيل المثال: من خلال خبرة سابقة أو شهادات).	• تقييم الشباب (فيما يخص الدبلومات والشهادات وغيرها). • تقييم السوق
	قيود على الجماعات المهمشة (عرقية، نوعية وغيرها)	• قد تترك العادات المحلية والمعايير الاجتماعية مجموعات معينة من الأشخاص من المشاركة في سوق العمل. • قد يحدث العزل المهني على أساس عرقي أو ديني.	• تقييم السوق و/ أو مسح أرباب العمل الحاليين (أفضلية التعيين) • تقارير حقوق الإنسان • الدراسات الإثنوولوجية

المصدر: مقتبس عن كينيث غهام وسانشين-جويرتو-فورتيلي (٢٠١٠).

الأهداف والتصميم: ما الذي نريد تحقيقه وكيف يمكننا تحقيقه؟

في ضوء المشكلات المحددة والمجموعات المستهدفة والحوافز المحددة أمام الفرص الاقتصادية الأفضل، يمكننا صياغة أهداف للبرنامج والاختيار من بين التدخلات الممكنة. حيث سيساعدنا التحديد الواضح لما نريد تحقيقه في التفكير في النتائج النهائية لبرنامجنا، والتواصل مع المتبرعين والمساهمين وإدارة التدخل ومتابعة وتقييم عملنا.

وضع هدف تطوير المشروع

تتمثل الخطوة الأولى في تحديد هدف تطوير مشروعنا^١. يمثل هدف تطوير مشروعنا ما نريد إنجازه والنتيجة المرجوة أو المخطط لها لتدخلنا. يمكن أن تساعدنا عدة مهام في وضع هدفنا الأولي مثل تحديد الجماعة المستهدفة بوضوح ومقدار التغييرات المتوقعة والفترة الزمنية (أنظر شكل ٢-٢). كلما كان الهدف ملموساً كلما كان من الأسهل تتبع التقدم حياله.

شكل ٢-٢ كيفية وضع أهداف المشروع

استناداً إلى مشكلات وقيود معينة تم تحديدها:

١. حدد النتيجة المتوقعة - ما الذي نتوقع تحقيقه وليس ما نريد القيام به.
٢. حدد الفئة المستهدفة.
٣. حدد اتجاه ومقدار التغييرات المتوقعة.
٤. حدد الفترة الزمنية التي ستظهر فيها التغييرات المتوقعة.
٥. تأكد أن الهدف قابل للقياس.
٦. تأكد أن الهدف يمكن أن يعزى إلى التدخل.



يوجد خطأ شائع عند تحديد هدف تطوير مشروعنا ألا وهو التركيز على ما سنفعله بدلاً من ما الذي ننوي تحقيقه (أنظر نقطة ١ في شكل ٢-٢). إذا كان السبب النهائي لتدخلنا هو تحسين الظروف المعيشية للشباب يجب عندئذ أن ينعكس ذلك على هدف مشروعنا. إن الوسيلة التي نحقق بها هذا الهدف - على سبيل المثال يعد توفير الدعم النفسي-الاجتماعي أو التدريب أو رأس المال المبدئي أو خدمات أخرى - هو «الكيفية» وليس الهدف الفعلي. مربع ٢-٣ يقيم ثلاثة أمثلة لهدف تطوير المشروع.

١ تستخدم المنظمات مصطلحات مختلفة لتسمية أهداف تطوير مشروعاتها، كهدف المشروع والهدف النهائي أو الغرض.

مثال أ	مثال ب	مثال ج
١٠٠٠ من الشباب في بيرو تم تدريبهم على مهارات الأعمال.	بحلول عام ٢٠١٥، تضاعف دخل ١٠٠٠ من الشباب المتسربين من المدرسة الذين تتراوح أعمارهم بين ١٨-٢٩ سنة في ليما ببيرو عن طريق (١) تعليمهم مهارات العمل و(٢) توفير رؤوس أموال مبدئية لهم	بحلول عام ٢٠١٥، انخفضت بطالة الشباب في ليما - بيرو بنسبة ١٠٪.

التقييم



هذا الهدف، على الرغم من كونه محدداً وقابلًا للقياس، يستحيل إسناده إلى مشروعنا. إن التدخل المفرد بنطاقه المحدود لن يتمكن من إحداث التغيير للمستوى المرتفع المرغوب حيث ستتأثر بطالة الشباب بعوامل متنوعة.



هذا المثال يلبي جميع متطلبات هدف المشروع الجيد فهو ملموس ويشير إلى تحسن قابل للقياس في الأحوال المعيشية للمجموعة المستهدفة.



مثال الهدف السابق يفتقر إلى عدة تفاصيل ضرورية (الفترة الزمنية، المجموعة المستهدفة تحديداً وغيرها) مما يجعلها عامة للغاية. علاوة على ذلك، فهو لا تشير إلى نتيجة متوقعة بل إلى خدمة سيتم توصيلها. إن حقيقة أنه سيتم تدريب الشباب لا تعني بالضرورة أنه سيحدث تحسن في وضعهم مثل زيادة المعرفة أو وضع التوظيف أو الدخل. يجب أن يذهب هدف المشروع إلى ما هو أبعد من ذلك.

- النتيجة المتوقعة
- الاتجاه والمقدار
- قابل للقياس
- قابل للإسناد

«الكيفية» (اختيارية)

مثال ب
بحلول عام ٢٠١٥، تضاعف دخل ١٠٠٠ من الشباب المتسربين من المدرسة الذين تتراوح أعمارهم بين ١٨-٢٩ سنة في ليما - بيرو عن طريق (١) تعليمهم مهارات العمل و(٢) توفير رؤوس أموال مبدئية لهم.

- الفترة الزمنية
- السكان المستهدفين

وضع الأهداف المؤسسية

بالإضافة إلى تحديد هدف تطوير المشروع، قد نهتم كذلك بتحديد الأهداف المؤسسية. ترتبط الأهداف المؤسسية بتدخلنا ولكنها قد لا تعود مباشرة على مجموعتنا المستهدفة الأولية. على سبيل المثال، يمكن أن تكون الأهداف المؤسسية داخلية بالنسبة إلى منظمنا. قد تكون الأهداف المؤسسية مثل تحديد الدروس للاستفادة من المشروع من ناحية التصميم والتنفيذ. قد تكون الأهداف المؤسسية متعلقة ببيئة المشروع مثل بناء الشراكات أو تعزيز الإرادة السياسية أو تحسين التزام المساهمين. فلكل هذه العوامل أهمية ويجب تحديدها منذ بداية المشروع.

تحديد التدخل

مع وجود هدف واضح في ذهن، يمكننا تحديد نطاق التدخل الذي سيقودنا إلى تحقيق أهدافنا المذكورة. من الطبيعي أن اختيار البرنامج يجب أن ينتج مباشرة عن الحواجز الخاصة المحددة في القسم السابق؛ أي أننا يجب أن نختار تدخل يتناول على نحو بين الأسباب الضمنية التي تعوق قدرات الشباب وفرصهم في تحقيق معيشة كريمة لأنفسهم وعائلاتهم.

وضع برامج قائمة على أدلة

من العناصر الحاسمة في تطوير التدخل هو مراجعة المعرفة القائمة عن بدائل البرنامج المتنوعة. على سبيل المثال، لمعالجة قيود بدء الأعمال للشباب فقد نرغب في تنفيذ برنامج يعمل على تطوير مشروعات الشباب ولكن ما الذي يجب أن يبدو عليه التدخل تحديداً؟ فرضاً أننا استطعنا تأكيد أن القيود المالية هي العقبة الرئيسية لبدء الأعمال. هل يجب أن يوفر البرنامج منح أو قروض؟ هل يجب أن يستهدف الشباب الأصغر أم الأكبر سناً؟ الأقل أم الأفضل تعليماً؟ وهل سيكون الدعم المالي كافياً أم أنه يجب أن يكون مصحوباً بخدمات دعم أخرى مثل التدريب وتقديم المشورة ودعم تطوير الأعمال؟

للإجابة عن هذه الأسئلة الصعبة سيستفيد مدراء البرنامج حتماً من الإطلاع على قاعدة الأدلة القائمة. في العديد من المرات كنا نتجه (أو المنظمات التي عملنا لديها) نحو تفضيل أنواعا بعينها من المشروعات بناء على نزعاتنا وخبرتنا السابقة إلا أنه لتطوير مشروعات عالية الجودة من المهم دراسة المعرفة النظرية والتجريبية عن عمل برامج سبل العيش للشباب. (يتضمن قسم المصادر في نهاية هذا الدليل مراجع لدوريات أكاديمية وقواعد بيانات وتقييمات أثر سابقة وجارية). فإذا أكدت الأدلة المتاحة ميولنا نحو تصميم معين، فيمكننا أن نؤسس قضية قوية لتصميم محدد. أما إذا، عوضاً عن ذلك، أشارت المعارف القائمة إلى وجود قيود كبيرة على التدخل فإن إدراج وتطبيق الدروس المستفادة في المبادرة الجديدة سيكون من شأنه توفير الوقت والمال.

يقدم جدول ٢-٢ أمثلة على التدخلات التي لديها سجل جيد بناء على تقييمات أثر سابقة أو بناء على بيانات متابعة إيجابية. وسوف يساعد الاستناد إلى مثل هذه البرامج في تصميم أفضل لهذه التدخلات بما يجعلها أكثر مصداقية. فلا يعني عدم وجود قاعدة أدلة أو وجود قاعدة محدودة أن التدخل المقترح محكوم عليه بالفشل. ففي الواقع، لن يوجد أي سجل للمناهج المبتكرة وذلك بسبب طبيعتها. وبالرغم من ذلك، فعند إجراء التدخلات التي تفتقر لقاعدة أدلة جيدة، يجب أن نكون دائماً على دراية بطبيعتها التجريبية وعدم التسليم بأن النتائج ستكون إيجابية. هنا يكون التقييم الحثيث هاماً بشكل خاص.

[معلومة إرشادية]

يعد سجل وظائف الشباب
(www.youth-employment-inventory.org)

مصدراً واحداً للتدخلات الجارية
والسابقة لتوظيف الشباب. حيث تتيح
قاعدة البيانات الديناميكية التصفح
والفرز حسب نوع التدخل والتقييم، بما
يمكن المستخدمين من البحث عن الأدلة
المتاحة عن نوع معين من المشروعات.

جدول ٢.٢ قائمة تدخلات قائمة على أدلة حسب القيود

القيود	تدخل مدعوم بدليل قوي	تدخل مدعوم بدليل مختلط
مهارات أساسية غير كافية	• معلومات عن قيمة التعليم	• برامج الفرصة الثانية للتعليم
عدم تطابق المهارات الفنية	• برامج تدريب إضافية / شاملة بالإضافة إلى معلومات المنافع الناتجة عن التخصصات الفنية	• التدريب في مواقع العمل
عدم تطابق المهارات السلوكية	لا يوجد	• التدريب على المهارات السلوكية / الحياتية
مهارات تنظيم مشروعات غير كافية	لا يوجد	• التدريب على تنظيم المشروعات
اقتصاد ذو معدل بطيء لنمو الوظائف	• إعانات للأجور أو إعانات للبرامج التدريبية	• برامج الخدمة العامة • الأشغال العامة كثيفة العمالة
تمييز رب العمل	• برامج التمييز الإيجابي	• إعانات لأرباب العمل الذين يعينوا المجموعات المستهدفة • الإشراف على الموظفين
نقص فرص الحصول على رأس المال أو رأس المال الطبيعي، أو رأس المال الاجتماعي	• برامج شاملة لتنظيم المشروعات	• تمويل المشاريع الصغيرة
المطابقة بين أرباب العمل وطالبي التوظيف	• خدمات التوظيف	• مشاركة المعلومات باستخدام التكنولوجيا
تمييز الكفاءات	لا يوجد	• شهادات المهارات • اعتماد مراكز التدريب
قيود على المجموعات المستبعدة (العرق، النوع، وغيرها)	• مشاركة المجموعة المستهدفة المهمشة في البرامج • تدريب غير تقليدي على المهارات • أماكن تدريب آمنة / أماكن عمل لمجموعات خاصة	• تعديل محتوى/تصميم البرنامج ليشمل حاجات المجموعات المهمشة

المصدر: مقتبس عن كينغهام وسانشيز-بيورتا وويرملي (٢٠١٠).

وبالنظر إلى التنوع الاقتصادي والاجتماعي والمؤسسي والإداري داخل وعبر البلدان، بالإضافة إلى الاحتياجات الخاصة للمجموعة المستهدفة، لن تكون كافة التدخلات المذكورة في جدول ٢-٢ بالضرورة ذات جدوى في سياق معين. فيجب تقييم ما إذا كانت الشروط المسبقة تتواجد في البلد المستهدفة أو سوق العمل المستهدف من عدمه، وإذا لم تتواجد، ما إذا كان من الممكن تعديل تصميم البرنامج لجعل التدخل ذا جدوى (كينغهام وسانشيز-بيورتا وويرملي ٢٠١٠).

العلاقة بين تصميم البرنامج والتقييم

أخيراً، من المهم أن نقر أن هناك صلات هامة بين تصميم البرنامج والتقييم. وكما رأينا في بند ١، حيث أن أحد الأدوار الرئيسية للتقييم هو دعم التعلم، وبالتالي، التخطيط. وبناء عليه، تعتمد فائدة وجدوى* التقييم بشكل كبير على جودة التصميم الأصلي للبرنامج. ويجب أن نضع النقاط التالية في الاعتبار:

- **التقييم لا يعرض عن سوء التصميم.** فلا يحل التقييم اللاحق محل التفكير المبكر. وأفضل ما يمكننا عمله من أجل برنامج ناجح هو تصميم مدروس للبرنامج يستند على البحث والخبرة الحالية.
- **ستعتمد استراتيجية التقييم على الفجوات المعرفية التي يتم تحديدها أثناء مرحلة التصميم.** تعد معرفة قاعدة الأدلة والتعرف على الفجوات المعرفية المحتملة عناصر هامة في اختيار الاستراتيجية الصحيحة للتقييم فيما بعد. على سبيل المثال، سوف تبرز أهمية تقييمات الأثر بصفة خاصة للبرامج المبتكرة وغير المختبرة التي تتيح فرصة ملء الفجوات في المعرفة على مستوى العالم.

[معلومة إرشادية]

بأخذ نظرة شمولية على تنمية الشباب، ينبغي فهم استراتيجيات تعزيز سبل العيش في سياق أوسع لما يحتاجه الشباب للانتقال إلى مرحلة الرشد بنجاح. فعلى سبيل المثال، يقدم معهد البحث إطار الأصول التنموية أربعين أصلاً داخلياً وخارجياً للشباب، والتي يمكن تقويتها من أجل تعزيز التنمية الإيجابية للشباب.

لمزيد من المعلومات، قم بزيارة:
<http://www.search-institute.org/developmental-assets>

- **التصميم الصحيح للبرنامج يمكن أن ييسر عملية التقييم.** بعض البرامج أسهل في التقييم من برامج أخرى. على سبيل المثال، إذا لم يتم التخطيط لتقييم الأثر أثناء مرحلة تصميم البرنامج، سيحدد ذلك الأدوات المتاحة للتقييم (انظر بند ٦). وبدورها، فإن عملية اختيار معايير استهداف واضحة وعادلة وشفافة مثل الاختيار العشوائي للمشاركين في برامج تجاوز عددهم فيها العدد المطلوب أو معدلات التأهيل، يمكنها بشكل بارز أن تيسر من عملية التقييم. وبناء عليه، إذا كان هناك عدة طرق مقبولة لتنفيذ برنامج معين، فإنه قد يكون من الحكمة التخطيط مسبقاً واختيار التصميم الذي يناسب التقييم أيضاً.

النقاط الرئيسية

١. حيث أن الفائدة التي تعود من المتابعة والتقييم تعتمد في النهاية على جودة التصميم الأصلي للبرنامج، فيجب أن نضمن معايير مرتفعة في تخطيط التدخلات الخاصة بنا وتصميمها.
٢. لغرض تصميم مشروعات عالية الجودة، يجب أن نفهم الشباب والسياق الذي يعيشون فيه. وهذا يتطلب تقييمات خاصة بالشباب والسوق تشمل عدة قطاعات، بشكل يستوعب التعقيدات في العوامل البيئية التي تؤثر على رفاهية الشباب وفرصهم.
٣. من الضروري تشخيص العناصر الحقيقية التي تعترض سبيل الشباب للحصول على وظيفة ودخل مناسبين. حيث إنه بدون معرفة ما الذي يحد فرصهم بالضبط، يستحيل تصميم تدخل يعالج القيود ذات الصلة.
٤. عند تصميم تدخل لتحقيق الهدف الإنمائي الوارد في المشروع، يجب الرجوع إلى الأدلة النظرية والتجريبية لزيادة احتمال النجاح وتجنب الأخطاء المكلفة.

دراسة حالة لصندوق العمل الاجتماعي لشمال أوغندا (NUSAF): مراجعة تصميم المشروع

تحليل المشكلة

يعتبر ارتفاع مستويات بطالة الشباب ونقص العمالة من المشكلات المستمرة التي تظهر على قمة أجندة السياسات لعدد من الحكومات في البلدان ذات الدخل المنخفض والمتوسطة. وينطبق هذا أيضاً على حكومة أوغندا، والتي تبحث عن طرق للتخفيف من احتمالات نشوب الصراعات في المستقبل في شمال البلاد.

التشخيص

في أفريقيا بصفة عامة، وفي شمال أوغندا تحديداً، لا يكاد يوجد فرص عمل في القطاع الرسمي لتلك الشعوب نظراً لعدم وجود شركات خاصة. وبالنظر إلى انعدام فرص العمل مع تدني مستويات المهارات ووجود عقبات لبدء أي عمل تجاري، قرر برنامج فرص الشباب التركيز على برنامج شامل لتنظيم المشروعات والذي سوف يقدم برامج تدريب على المهارات المهنية ومنح نقدية وخدمات دعم أخرى.

(...)

دراسة حالة لصندوق العمل الاجتماعي لشمال أوغندا (NUSAF): مراجعة تصميم المشروع

الأهداف والتصميم

كان لدى برنامج فرص الشباب هدفاً رئيسياً: بحلول عام ٢٠١٠، حيث سعى البرنامج إلى زيادة التوظيف لعدد ٨,٠٠٠ شاب على الأقل، ممن تتراوح أعمارهم بين ١٥ - ٣٥ عاماً في شمال أوغندا وذلك بالترويج لشركات تعتمد على المهارات وبناء قدرات مراكز التدريب (لم يتم تحديد الحجم المطلوب من تأثيرات العمل).

بالإضافة إلى الهدف الرئيسي السابق، تناول البرنامج عدداً من الأهداف الثانوية، مثل تحسين التواصل الاجتماعي للشباب في مجتمعاتهم وتخفيف الألم النفسي الناتج عن الخلافات الحالية. وبشكل أعم، كان يهدف البرنامج إلى المساهمة في رفع رفاهية الشباب وأسرهم المعيشية بشكل شامل، وتحسين الصحة ومستوى الحياة وتوفير نمو اقتصادي مستدام، وبالتالي، تقليل إمكانية نشوب صراعات في المستقبل في شمال أوغندا.

المصدر: باتلمان ومارتينيز وفيالا (٢٠٠٩).

قراءات رئيسية

بيدويل، كيه، غالريث، سي، وآخرون، ٢٠٠٨.

Market Assessment Toolkit for Vocational Training Providers and Youth

نيويورك: اللجنة النسائية المعنية باللاجئين من النساء والأطفال وكلية كولومبيا الجامعية للشؤون الدولية والعامة.

http://www.womensrefugeecommission.org/docs/ug_ysl_toolkit.pdf

كينغهام، دبليو، سانشيرز - بيورتا، إم. إل.، وويرملي، أيه. ٢٠١٠.

Active Labor Market Policies for Youth: A Framework to Guide Youth Employment Interventions

واشنطن، دي سي: البنك الدولي.

http://siteresources.worldbank.org/INTLM/214578-1103128720951/22795057/EPPNoteNo16_Eng.pdf

صندوق الأمم المتحدة لتنمية رأس المال. ٢٠١١.

Listening to Youth: Market Research to Design Financial and Non-Financial Services for Youth in Sub-Saharan Africa

نيويورك: صندوق الأمم المتحدة لتنمية رأس المال

<http://www.uncdf.org/english/microfinance/uploads/other/Listening%20to%20Youth-YouthStart%20Market%20Research.pdf> UNCDF.

اللجنة النسائية المعنية باللاجئين. ٢٠٠٩.

Building Livelihoods: A Field Manual for Practitioners in Humanitarian Settings

نيويورك: اللجنة النسائية المعنية باللاجئين WRC

http://www.womensrefugeecommission.org/docs/livelihoods_manual.pdf

بند ٣: إرساء نظام متابعة

ما يتم قياسه يتم إنجازه.

— توم بيترز

من المستحيل القيام بتقييم جيد دون وجود نظام متابعة جيد. وعلاوة على ذلك، فإن تصميم نظام متابعة جيد من شأنه أن يحسن على الأغلب من جودة تصميم مشروعنا وسيسهل إدارة المشروع. ويلخص هذا البند الخطوات الرئيسية التي ينبغي إتباعها لبناء نظام متابعة في أي مشروع، سواء كان سيتم إجراء تقييم أم لا. وكما سنرى، يجب أن توجد أدوات المتابعة التالية، كحد أدنى، في كل مشروع:

- سلسلة نتائج.
- إطار منطقي.
- نهج عملي من أجل جمع المعلومات وتحليلها وتطبيق النتائج.

لماذا نحتاج إلى نظام متابعة؟

توفر المتابعة معلومات داخلية وخارجية بشكل متواصل وذلك من أجل إطلاع مدراء البرنامج حول التطورات المخططة والتطورات الفعلية. وعندما يتم متابعة مخالفات أو أوجه قصور، يمكن تصحيحها في الوقت المناسب. وتتضمن المتابعة جمع وتحليل البيانات من أجل التحقق من استخدام الموارد حسب الغاية المنشودة، وأنه يتم تنفيذ الأنشطة بموجب الخطة وتسليم المنتجات والخدمات المتوقعة، وأن يتم الوصول للمستفيدين المستهدفين (سافيدوف وليفين وبيردسال ٢٠٠٦). ويجب أن تكون المتابعة الفعالة جزءاً من كل مشروع، بغض النظر عما إذا كان سيتم تقييم المشروع أم لا.

توفر المتابعة أيضاً الأساس لتقييم تدخل ما. وفي الواقع، فإنه من الصعب إجراء تقييم جيد دون توفر معلومات ملائمة حول التنفيذ الفعلي. ففي حالة عدم توفر معلومات موثوقة حول تقدم التنفيذ وجودته، عندها سيكون أي تقييم عرضة لخطر إساءة تفسير أسباب نجاح أو فشل المشروع.

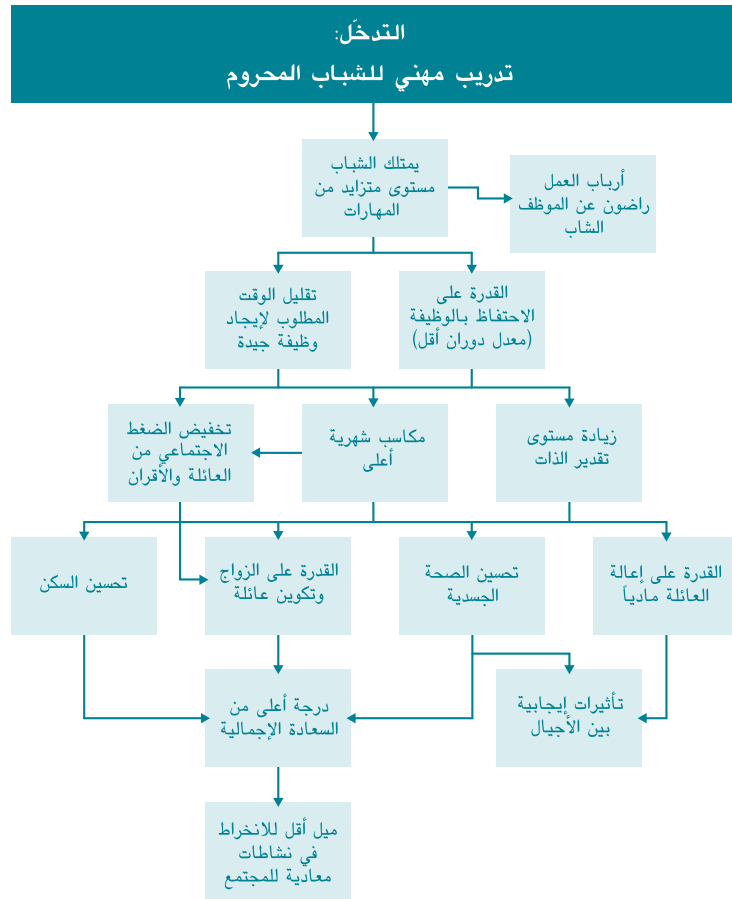
تتمثل تحديات إجراء متابعة تدخل ما في:

- تحديد المنطق وراء التدخل، والذي يتضمن تحديد أهداف تتخطى الهدف التنموي للمشروع وتغطي كل مستويات التنفيذ والنتائج.
- تحديد المؤشرات الرئيسية وآليات جمع البيانات والافتراضات التي يمكن استخدامها لمراقبة التطور الذي يتم احرازه مقابل هذه الأهداف.
- إرساء نظام للمتابعة ورفع التقارير لتتبع التقدم المحرز من أجل تحقيق الأهداف وإفادة مدراء المشروع وغيرهم من أصحاب المصالح.

تحديد المنطق وراء التدخل

العلاقة بين تصميم المشروع ونظرية المشروع

إن نظرية التغيير موجودة في تصميم أي برنامج. وكما تم نقاشه في بند ٢، يوجد عادة توقع أن مشروعاً ما سيساعد في تحسين الظروف المعيشية للمجموعة المستهدفة عبر معالجة مجموعة محددة من الحواجز والعوائق التي يواجهها هؤلاء الشباب. وهذا يعني أنه يوجد لدينا مجموعة من الافتراضات حول كيف ستعزز أنشطة مشروع معينة التغيير الإيجابي ولماذا. لماذا نعتقد أن تدريب الشباب سوف يؤدي إلى نتائج أفضل في سوق العمل؟ لماذا نعتقد أن دعم مشاريع الشباب سيؤدي إلى الحد من الفقر؟ من أجل التأكيد على ملائمة تدخلنا، يجب أن تكون النظرية وراء هذا التدخل واضحة (انظر شكل ٣-١).



يجب على الممارسين توضيح نظرية تغيير لكل تدخل. ومثالها، يتم تطوير نظرية التغيير في بداية مرحلة تصميم المشروع، عندما يمكن جمع جميع أصحاب المصالح للموافقة على رؤية مشتركة للمشروع، وعلى غاياته المحددة، والخطوات اللازمة لتحقيق هذه الغايات (غير تالر وآخرون ٢٠١١). وحسب تايلور-بويل (٢٠٠٥)، يساعد استخدام نظرية تغيير مدير المشروع والمقيم على حد سواء من خلال:

- زيادة فهم البرنامج وتوفير لغة مشتركة.
- المساعدة في التفريق بين «ما نقوم به» و«ما نريد تحقيقه».
- تحسين التخطيط والإدارة.
- تحديد المتغيرات المهمة التي يجب قياسها.
- توفير أساس لإجراء تقييمات متعمقة.

تحويل النظرية إلى سلسلة نتائج

عملياً، يمكن تطبيق نظرية التغيير بعدة طرق مختلفة. وتتضمن التطبيقات الشائعة: النماذج المنطقية أو الأطر المنطقية أو نماذج النتائج أو سلاسل النتائج. والفكرة هي دائماً نفسها: منح أصحاب المصالح «ملخص منطقي ومعقول» عن كيف يمكن للتدخل المخطط له أن يؤدي إلى النتائج المنشودة (غير تالر وآخرون ٢٠١١، صفحة ٢٤). أنظر شكل ٣-٢. وكنتيجة، فإنها تقدم تسلسلاً من الأحداث التي تربط العناصر الواقعة تحت المسؤولية المباشرة للمشروع (الموارد المستخدمة والأنشطة المنفذة، والمنتجات والخدمات المقدمة) مع النتائج المتوقعة وغايات البرنامج رفيعة المستوى.

شكل ٣-٢ مكونات سلسلة النتائج وأمثلة على ذلك



* مستوى هدف تطوير المشروع

تتضمن عملية التنفيذ المخطط الخاصة بنا الفئات التالية ويكون مدير البرنامج مسؤولاً مسؤولية مباشرة عن:

- **المدخلات** — الموارد المتوفرة للمشروع، وتتضمن الميزانية والعاملين والشركاء والمعدات.
- **الأنشطة** — إجراءات البرنامج وعملياته وتقنياته وأدواته وأحدثاته والتكنولوجيا الخاصة به. وتستخدم صيغة الفعل لوصف هذه الأنشطة (وفر، يسر، وصل، نظم، وغيرها).
- **المخرجات** — المنتجات والخدمات المقدمة والتي تخضع مباشرة لرقابة المنظمة المنفذة. وهي تشير إلى إذا ما تم تسليم البرنامج حسب ما هو مقصود. ويتم عادة التعبير عن المخرجات بإجراءات منجزة (تم التدريب، تمت المشاركة، تم الاستخدام، تم التمويل، وغيرها).

[تعريف]

سلسلة النتائج هي سلسلة من الموارد والأنشطة والخدمات التي يتم تقديمها والتي يتوقع أن تؤثر على الآثار المباشرة وطويلة الأمد على مجتمع الدراسة المستهدف.

[معلومة إرشادية]

على الرغم أنه ليس ضرورة مطلقة، إلا أنه غالباً ما تكون فكرة تضمين أهدافكم المؤسساتية والأنشطة الأساسية في سلسلة النتائج فكرة جيدة.

تصف نتائجنا المنشودة جميع آثار البرنامج المرغوبة حسب الفئات التالية:

- **النتائج** — الآثار على المدى القصير والمتوسط (عادة ضمن عدة أشهر وحتى سنتين) على المجتمع المستفيد من الدراسة والناشئة عن مخرجات المشروع. ويمكن أن تتضمن هذه النتائج تغيرات في المواقف والمعارف والمهارات، والتي يمكن أن تكون آثاراً مباشرة نسبياً، كما يمكن أن تكون تغيرات في السلوكيات والأوضاع وما شابه والتي يمكن أن تستغرق وقتاً أطول. وينبغي أن تكون النتائج الرئيسية المستهدفة هي التي تم تحديدها في هدف تطوير المشروع. ويتم عادة التعبير عن النتائج على مستوى فردي وتشير هذه النتائج إلى تغيير ملحوظ (زيادة، تحسن، انخفاض، وغيرها).
- **نتائج ذات مستوى عال** — أهداف المشروع طويلة الأمد والتي تتعلق عادة بالمعايير المعيشية الإجمالية. ويمكن أن تتأثر هذه النتائج بعوامل متنوعة ولا تقع عادة تحت السيطرة الكاملة للبرنامج. وغالباً ما يسمى هذا المستوى من سلسلة النتائج بـ «الآثار»، ولكننا نفضل عبارة «نتائج ذات مستوى عال» لتجنب الالتباس حول المعنى المحدد لكلمة «أثر» في إطار تقييم الأثر (انظر بند ٥).

إنشاء سلسلة نتائج

تحديد مستوى الملاحظة

قد نرغب في ملاحظة عوامل أخرى غير تلك المرتبطة بالشباب كأفراد، وذلك فيما يتعلق بالتنفيذ والنتائج على حد سواء. وفي الواقع، قد نكون أيضاً مهتمين بالمخرجات أو النتائج على مستوى الأسرة أو على مستوى المجموعة أو المرفق (المدارس، مراكز التدريب المهني) أو على مستوى القرية/المجتمع.

أخذ التنوع الممكن في النتائج المحتملة

يمكن أن تؤثر تدخلات سبل عيش الشباب على العديد من النتائج، بما في ذلك النتائج المتعلقة مباشرة بالفرص الاقتصادية وسوق العمل ولكنها تتعداها بكثير. وحسب التدخل، قد يكون من المفيد استهداف وقياس مجموعة متنوعة من النتائج إذا لاءمت هذه النتائج المنطق وراء المشروع والأهداف. وتتضمن فئات النتائج الشائعة ما يلي:

- **النمو النفسي الاجتماعي** — قياسات لعقل الشاب ومشاعره ومستوى نضجه. ويمكن أن تتعلق النتائج بتقدير الذات أو الهوية أو الثقة أو العزلة أو الصحة النفسية والاجتماعية.
- **المهارات** — مستويات المعارف الأساسية في القراءة والكتابة والحساب، والكفاءات الفنية في مهنة محددة (الحرفة والميكانيكا والمحاسبة وخدمات العملاء) والمهارات الحياتية (الاتصال والتواصل والعمل الجماعي والتفكير الناقد والإدارة الذاتية)، ومهارات تنظيم المشاريع (الإبداع والمهارات التجارية).
- **التوظيف وسوق العمل** — استخدام المستفيدين للوقت (بين المدرسة والعمل مدفوع الأجر والعمل الحر والبطالة والعمل المؤقت)، وخصائص الوظيفة (نمط رب العمل أو العمل وعدد ساعات العمل والدخل)، وخصائص العمل التجاري (الأرباح وعدد الموظفين واستمرار العمل التجاري ومعدل سداد القروض).
- **استخدام الخدمات المالية** — إمكانية حصول المستفيدين على الخدمات المالية والسلوكيات المتعلقة بالتعاملات البنكية وتوفير المال وإدارة الديون والميزانيات والرفاه المالية العامة.

- **السلوكيات الخطرة** — المواقف والسلوكيات المتعلقة بالكحول وتعاطي التبغ والعقاقير والصحة الإنجابية (مثل: ممارسة الجنس دون وقاية وفيرس نقص المناعة المكتسبة/ الإيدز) والجريمة والعنف.
- **تكوين العائلة** — المواقف والسلوكيات المتعلقة بسن الزواج والعدد المرغوب والفعلية للأطفال.
- **المواطنة** — تفضيلات الشباب وأفعالهم فيما يتعلق بالتصويت في الانتخابات المحلية والوطنية، والمشاركة في المجتمع (مثل المشاركة عبر العضوية في ناد أو التطوع) وتولي مناصب قيادية.
- **الاستثمارات في رأس المال البشري** — التغيرات في الوضع التعليمي (عاد إلى المدرسة أو يرغب بالعودة إلى المدرسة) ومبلغ المال المنفق على التعليم أو الصحة (على النفس أو الغير)، والمساهمات بين الأجيال (مثل: التطعيم ومتابعة نمو أطفاله).
- **أمور أخرى** — قد تتعلق النتائج الإضافية بأنماط الاستهلاك والتغذية وإيجاد الأصول والقدرة على التنقل والهجرة بالإضافة إلى العلاقات العائلية والعلاقات مع المجتمع.

أخذ النتائج غير المقصودة في الاعتبار

يعكس هدف مشروعنا النتيجة الرئيسية المرغوبة للتدخل إلا أن مشروعات التطوير تتسم بالتعقيد وقد يكون لتدخلنا تأثيرات غير مقصودة. قد تكون بعض هذه التأثيرات غير المقصودة متوقعة في حين يكون البعض الآخر غير متوقع ومفاجئ. قد تكون كل من النتائج المتوقعة وغير المتوقعة إيجابية أو سلبية (أنظر شكل ٣-٣). من المهم تضمين هذه النتائج المحتملة (أنظر الفئات الرئيسية أعلاه) في سلسلة النتائج وتصنيفها طبقاً لذلك بغرض الوقوع على المنطق الكامل وراء التدخل بشكل واقعي وتوفير أسس متابعة جميع الآليات المطبقة.

شكل ٣-٣ النتائج المقصودة في مقابل النتائج غير المقصودة



المصدر: مقتبس من هيمبل (٢٠٠٦).

على سبيل المثال قد تظهر امتدادات غير مباشرة من تدخلنا لأن المشاركين ينقلون معرفتهم إلى أفراد العائلة أو المجتمع الذين قد يستفيدوا بدورهم بشكل غير مباشر أيضاً، ونحن نرغب بالتأكيد في الوقوع على هذا التأثير. من ناحية أخرى، قد تظهر تأثيرات سلبية غير متوقعة: في مشروع تنظيم المشروعات، على سبيل المثال، قد يجد بعض الشباب أنفسهم غارقين في الديون لأن أعمالهم لم تستمر، في حالات أخرى حيث يحقق الشباب دخلاً أعلى بفضل تدخلنا، قد يكف أفراد العائلة عن العمل أو قد يستخدمون الدخل الإضافي في الإدمان في سلوكيات غير صحية مثل تعاطي المشروبات الكحولية والتبغ. ونكرر أننا نرغب في معرفة ما إذا كانت هذه التأثيرات ستقع. وكثيراً ما ساعد إجراء بحوث عن مشروعات مماثلة في تحديد نطاق النتائج المحتملة الإيجابية والسلبية.

في الأعوام الأولى من الألفية الثالثة، نفذ مجلس السكان مع مؤسسة إنقاذ الأطفال مشروع إشراق في الأماكن الريفية في صعيد مصر بتأسيس مساحات صديقة للفتيات لنقل مهارات الحياة لهن وبناء شبكات اجتماعية وتعزيز القيادة والثقة بالنفس. انتهى الأمر بأن ذهبت فوائد البرنامج لما هو أبعد من الفتيات المراهقات المتسربات من المدارس المستهدفات وامتدت إلى والدي المشاركات. نقلت الفتيات المعلومات التي حصلن عليها من فضولهن الدراسية إلى أمهاتهن وتضمنت هذه المعلومات معلومات صحية. بالإضافة إلى ذلك، فإن ملاحظتهن لمشاركة بناتهن في الحياة العامة كان له أثرًا قوياً على تصورات الأمهات عن وضعهن الخاص في المجال العام كما أنه بفضل اشتراك بناتهن في إشراق أدركت الأمهات أن لديهن الحق أيضاً في الحصول على الخدمات العامة.

المصدر: برادي وسالم وزباني (٢٠٠٧).

[معلومة إرشادية]

إذا طغت مخاطر متابعة النتائج غير المقصودة على إطار النتائج فقد تختار فرق المشروع التركيز على متابعة النتائج المقصودة واستخدام التقييمات لتحديد درجة النتائج غير المقصودة.

تجنب الأنشطة أو المخرجات المكررة

عند تطوير سلسلة النتائج الخاصة بنا، قد نحدد الأنشطة التي ليس لها علاقة وثيقة بالهدف الأساسي لمشروعنا. من أجل مشروع محدد جيداً وفعال يجب إغفال الأنشطة والمخرجات من هذا القبيل والتي تكون غير حاسمة.

تحديد المؤشرات الرئيسية وأدوات جمع البيانات والافتراضات

بمجرد أن يصبح لدينا سلسلة نتائج، كيف يمكننا أن نعرف إذا ما قد تم التخطيط له يتحقق بالفعل؟ من أكبر التحديات الكامنة في تطوير نظام متابعة هو تحديد طبيعة المعلومات التي تعكس مدى تحقيقنا لأهدافنا. نحن نحاول الآن تحديد المؤشرات المناسبة وأدوات جمع البيانات والافتراضات لكل مستوى من الأهداف بداية من المدخلات وحتى النتائج ذات المستوى العالي. يوفر الإطار المنطقي مصفوفة مفيدة لمعرفة جميع هذه العناصر (أنظر جدول ٣-١).

الخطوة ١: تحديد المؤشرات

المؤشرات التي تجيب عن سؤال «كيف سأعرف؟» وهذه المؤشرات هي

- جوانب رئيسية (أو إنابات) للعنصر الذي نريد قياسه برغم أنها قد لا تكون بالضرورة مثله بالكامل.

- علامات ملموسة على وجود شيء قد تم فعله أو شيء قد تم تحقيقه؛ وهي الوسائل التي نختارها كعلامات تدل على نجاحنا (شابيرو ٢٠٠٣).

تعد المؤشرات عنصراً حيوياً في نظام المتابعة لأنها تقود عملية جمع كافة البيانات التالية وتحليلها ورفع تقارير بها. فبدون وجود مجموعة واضحة من المؤشرات تفقد أنشطة المتابعة أو التقييم قدرتها على مقارنة التقدم الفعلي للبرنامج بما كان مخططاً ومتفقاً عليه (غوسباريني وآخرون ٢٠٠٤).

الخطوة ١			الخطوة ٢		الخطوة ٣	
الأنشطة	المخرجات	الأهداف	المؤشرات والغايات	مصدر المعلومات	التكرار	الطرف المسؤول
نتائج ذات مستوى عال	- خفض بطالة الشباب. - دخل أعلى للأسرة المعيشية.	(%) معدل البطالة المحلية. - دخل أعلى للأسرة المعيشية (بالدولار).	- إحصائيات التوظيف (على مستوى الوزارة، المدينة). - مسح للأسر.	سنوياً	فريق البرنامج	مهارات جديدة مطلوبة يظهر مردودها في سوق العمل.
النتائج	- تنفيذ المنهج على نطاق واسع. - فهم أفضل لآليات الأعمال. - تحسين المهارات الحياتية. - تحسين فرص التوظيف. - زيادة الاهتمام بالتعليم العالي.	- خلال ستة أشهر من إكمال البرنامج: ٥٠٠ مدرسة تستخدم منهجاً جديداً. ٥٠٪ الزيادة في إجابات صحيحة في الاختبار اليومي لمعرفة الأعمال. - نسبة ٧٠٪ من الطلاب راضين عن المنهج الجديد. - تحسن تصورات المدرسين والآباء عن المهارات الحياتية بنسبة ٣٠٪. - انخفاض الوقت المستغرق في البحث عن وظيفة بنسبة ٥٠٪ وزيادة رضا رب العمل بنسبة ٣٠٪. - زيادة ٥٪ في نسبة الالتحاق بالجامعة.	- مقابلة مع سلطة التعليم الرسمية. - نتائج الاختبارات المدرسية. - مجموعة نقاش يومية/مركزة مع المدرسين والآباء. - مسح رب العمل. - إحصائيات الالتحاق بالمدارس على المستوى الإقليمي.	نصف سنوي	- البرنامج - الفريق (المقابلات، جمع البيانات) - المستشار (المسح، حلقات النقاش)	- المنهج الذي قبلته سلطات المدارس المحلية. - جودة التدريس. - يمكن للشباب حضور المدرسة بانتظام.
المخرجات	- إجازة المنهج الجديد. - تم تدريب المدرسين. - تم تدريب الشباب على مهارات الأعمال.	- إنهاء البرنامج: - اعتماد منهج جديد من الوزارة. - تم تدريب ٥٠٠ مدرس. - أكمل ١٠,٠٠٠ شاب التدريب.	بيانات البرنامج.	نصف شهري	فريق البرنامج	- رغبة المدرسين في الحصول على التدريب. - يمكن للشباب حضور التدريب.
...	...	...	...	...	...	...

ملاحظة: حتى نكون عمليين قمنا بحذف الأنشطة وفئات المدخلات التي يتم تضمينها عادة في الإطار المنطقي.

اختيار مؤشرات لجميع مستويات سلسلة النتائج

حتى عندما ينصب تركيزنا على نتائج التدخل، من المهم متابعة مؤشرات على مستوى التنفيذ حتى تتمكن من تحديد ما إذا كان المشروع قد وصل إلى المستفيدين المقصودين وما إذا كان قد تم تنفيذه كما كان مقصوداً. وبدون هذه المؤشرات في سلسلة النتائج بأكملها، لن يحدد التقييم سوى ما إذا تم تحقيق النتائج المتوقعة، ولكنه لن يكون قادراً على إيجاد علاقة بين مستوى نجاح البرنامج وجوده تنفيذه. يوضح جدول ٣-٢ أمثلة على هذه المؤشرات على سلسلة النتائج بأكملها.

جدول ٣-٢ أمثلة على المؤشرات

الفئة	العينة المستهدفة	مثال على المؤشرات
المدخلات	اثنان من المدربين وتسهيلات في حدود ميزانية قدرها ١٠,٠٠٠ دولار أمريكي.	- اثنان من المدربين المهرة والمجهزين والموزعين. - تكلفة البرنامج بالدولار الأمريكي في حدود الميزانية المطلوبة.
النشاط	تقديم تدريب على مهارات الحياتية للشباب (٢٠ ساعة).	- عدد ساعات التدريب المقدمة. - عدد الشباب المشارك حسب السن والنوع ومستوى التعليم. - التاريخ الذي تم تقديم التدريب فيه.
المخرجات	مشاركة عدد ١٠٠ شاب في التدريب.	عدد الشباب الذين استكملوا التدريب (حسب السن والنوع ومستوى التعليم).
النتائج	زيادة المعرفة بالإتصال والتواصل الفاعل.	بنهاية البرنامج: - قياس عدد ونسبة الشباب القادر على التعبير عن الأفكار بوضوح مقابل بطاقة درجات الاختبار المحددة مسبقاً. - قياس عدد ونسبة الشباب الذين أظهروا تحسن في مهارات الإتصال والتواصل الشفوية وغير الشفوية مقابل بطاقة درجات الاختبار المحددة مسبقاً. - عدد ونسبة الشباب الذين عبروا عن الإحساس بالراحة عند تعاملهم مع أرباب العمل.
نتائج ذات مستوى عال	زيادة دخل الأسرة.	بحلول عام ٢٠١٥، زاد متوسط دخل الأسرة الشهري بنسبة ٢٠٪ مقارنة بنقطة الأساس.

يتطلب تحديد مؤشرات نتائج جيدة اهتماماً خاصاً. كما ذكر أعلاه، فنتائج التدخلات المتعلقة بسبل عيش الشباب يمكن أن تكون متنوعة للغاية ولكنها لا تقتصر على نتائج سوق العمل، ولذلك نحتاج إلى اختيار المؤشرات (النمو النفسي الاجتماعي والمهارات والتوظيف وغيرها) من بين جميع المجالات المناسبة. تعتمد المجالات الأدق التي سيتم قياسها على هدف التدخل وتركيزه وأهداف التعلم التي سيتم تناولها. فعلى سبيل المثال، قد يكون من المفيد، بالنسبة لتدخل مهارات الحياتية، قياس المهارات ونتائج سوق العمل والسلوكيات الخطرة في حين أن مشروع دعم إيجاد الوظائف، بدلاً من ذلك، قد يركز بشكل كامل على نتائج سوق العمل.

تحديد المؤشرات

اجتذاب أصحاب مصالح آخرين. يمكن أن يؤدي اختيار مؤشرات دون اشتراك مناسب من أصحاب المصالح من الداخل والخارج إلى عدم شعورهم بملكية وانتماء (كيوسك وريست ٢٠٠٤). يتعين التعاون مع الشركاء وأصحاب المصالح الأساسيين في المجتمع للوصول إلى مجموعة متفق عليها فيما بينهم من الغايات والأهداف ومؤشرات الأداء للبرنامج.

اختيار العدد الصحيح من المؤشرات. حيث إن المؤشرات ما هي إلا إنبات، فمن الشائع تحديد عدة مؤشرات لكل عنصر في سلسلة النتائج خاصة فيما يتعلق بالنتائج أو النتائج عالية المستوى إلا أن اختيار عدد زائد من المؤشرات سيعقد، بغير داع، نظام المتابعة الخاص بنا وسيزيد من أعباء جمع البيانات والتحليل ورفع التقارير. من المهم تحديد مؤشر واحد إلى ثلاثة مؤشرات تعكس بشكل أفضل كل عنصر في سلسلة النتائج.

[مصدر إلكتروني]

مؤشرات مخرجات ونتائج مختارة
عنوان الانترنت

احترام معايير الجودة. حتى مع عدم وجود مبادئ مطلقة لما يجعل من مؤشر ما مؤشراً جيداً، إلا أن خصائص سمات شذوذه الماثورة يمكن أن تكون مفيدة (غيرتير وآخرون ٢٠١١ ص. ٢٧).

مؤشرات SMART هي

- محددة، لقياس المعلومات المطلوبة عن كنب قدر الإمكان،
- قابلة للقياس، للتأكد من أن المعلومات يمكن الحصول عليها بسهولة،
- قابلة للإسناد، لضمان أن أي قياس سيكون مرتبطاً بمجهودات المشروع،
- واقعية، لضمان أنه يمكن الحصول على البيانات في وقت مناسب على فترات متكررة معقولة وبتكلفة معقولة، و
- تستهدف المجتمع المعني.

سيحدد اختيارنا من المؤشرات جزئياً بقدرتنا على جمع البيانات عنها. إذا لم يكن المؤشر قابلاً للقياس أو إذا لم تكن المعلومات متاحة فلن يتمكن عندئذ من خدمة غرضه في إظهار تقدم أهدافنا. إذا لم نكن قادرين على جمع بيانات لمؤشر قمنا باختياره فیتعين علينا أن نستبدله.

وضع نقطة أساس. تدلنا نقطة الأساس عن قيمة مؤشر ما في بداية أو - مثالياً - مباشرة قبل فترة التنفيذ. معرفة قيمة نقطة الأساس لمؤشراتنا تمكننا من تحديد أهداف واقعية وكذلك بتتبع التقدم المستقبلي مقابل الوضع الأولي. على سبيل المثال، إذا أردنا متابعة دخل المشاركين على مدار الوقت، فقد تدلنا البيانات في نماذج التسجيل الخاصة ببرنامنا على أن متوسط الدخل الشهري للمشاركين وقت اشتراكهم في البرنامج كان ١٠٠ دولار. هذه هي قيمة نقطة الأساس لدينا والتي يمكن أن تستخدم كمقارنة لكيفية نمو الدخل أثناء تدخلنا وبعده.

تحديد الغايات. إذا لم تكن المؤشرات محددة من ناحية الإطار الزمني والكم والجودة، فلن نستطيع أن نكون متأكدين بشكل كامل من أننا على المسار الصحيح وأنها سنصل إلى أهدافنا (كولي ١٩٨٩). على سبيل المثال، إذا كانت النتيجة المرجوة هي زيادة دخل الأسرة المعيشية فقد يكون مؤشرنا هو الأجور المكتسبة شهرياً بالدولار الأمريكي ثم يمكن أن نتحدد الغاية بحيث تكون زيادة بنسبة ثلاثين بالمائة (الكم) من التوظيف بالقطاع الرسمي (الجودة) في خلال ثلاث سنوات (الإطار الزمني). يجب ألا يكون لكل مؤشر أكثر من غاية واحدة لكل فترة زمنية معينة. إذا كان وضع غايات رقمية محددة عشوائي، فيمكن أن يتم التعبير عن الغايات كنطاق أو مدى (من إلى).

التأكد من الاتساق. على الرغم من عدم إمكانية ذلك في جميع الأوقات، إلا أنه كي نضمن اتساق المتابعة على مدى الوقت، فيجب أن نبذل جهداً لتذكر المؤشرات التي تم الاتفاق عليها قبل بداية المشروع. لذلك، فمن الشائع إضافة مؤشرات جديدة وإغفال المؤشرات القديمة عند تعديلنا لتصميم البرنامج أو تحسين نظام المتابعة. إلا أنه من الضروري تذكّر الأهداف الأصلية للمشروع. يجب أن تكون المتابعة والتقييم صادقين. إذا وجدنا أن مشروعنا لن يحقق هدفه الأصلي وأنه سيحقق هدفاً آخر بدلاً من ذلك (والذي قد يكون هدف ذا قيمة أعلى)، يجب أن نعترف بذلك عند رفعنا للتقارير. يجب عدم تغيير المؤشرات التي تم قبولها في بداية التدخل إلا إذا وجدت محددات موضوعية لتبرير التغيير.

يقدم جدول ٣-٣ أمثلة على مؤشرات التدخلات المتعلقة بسبل عيش الشباب على جميع مستويات سلسلة النتائج. في بعض الأحيان، يمكن استخدام مؤشرات محددة مسبقاً إلا أنه من المهم دراسة مواءمتها للمشروع المحدد. قد يلزمها بعض التعديل بحيث تتلاءم أو تكمل المؤشرات الأخرى الأكثر مواءمة محلياً.

[معلومة إرشادية]

عادة ما تكون تجربة المؤشرات أثناء المراحل المبكرة من التدخل قبل وضعها كجزء كامل لنظام المتابعة فكرة جيدة. حيث سيبرز ذلك كيفية عملها وما إذا كانت تجمع كافة المعلومات التي يهتم بها مدير المشروع وغيره من أصحاب المصالح.

النتاج الذي سيتم قياسه: تحسين فرص التوظيف للشباب - في سن ١٨ - ٢٤

مؤشر سبي: سيجد الشباب وظائف بسهولة أكبر مما كان بإمكانهم قبل التدخل

مؤشر جيد: عدد ونسبة الشباب - في سن ١٨ - ٢٤ الذين لديهم عرضين على الأقل لوظيفتين براتب أعلى من الحد الأدنى للأجور في مجال تدريبهم في خلال ثلاثة أشهر من استكمال البرنامج (مقارنة بعدم وجود عروض وظائف على الإطلاق من قبل).

جدول ٣-٣ أمثلة على المؤشرات الخاصة بمشاريع سبل عيش الشباب

نوع المشروع	المداخلات	الأنشطة	المخرجات	النتائج	نتائج المستوى العالي
التدريب وتنمية المهارات	• تخصيص الميزانية (بالدولار والنققات (الأمريكي) • مبلغ وحصة الأموال المطلوبة التي تم جمعها • عدد العاملين بالبرنامج حسب المستوى. • عدد الميسرين المحليين المتعاقد معهم • عدد المتابعات المحلية التي تقدم مساهمات عينية.	• عدد ورش العمل التي تم تنظيمها. • عدد ساعات التدريب. • عدد الشباب الذين تم اختيارهم/تسجيلهم. • عدد أرباب العمل الذين يقدمون فرص تدريب ميداني. • عدد فرص التدريب الميداني.	• عدد ونسبة الشباب الذين يحضرون على الأقل ٨٠٪ من التدريب. • عدد الشهادات التي منحت. • عدد الشباب الذين تم توفير فرص تدريب ميداني لهم. • متوسط فترة التدريب الميداني الذي تم إكماله (بالأسابيع).	• عدد ونسبة الشباب الراغبين عن البرنامج. • عدد ونسبة الشباب الذين يقدمون تقارير تفيد بوجود تصبن في قدرتهم على التفكير النقدي وعلى حل المشكلات. • عدد ونسبة الشباب يتلقون عروض وظائف بعد إتمام التدريب الميداني. • عدد أرباب العمل المحليين الذين يقدمون فرص عمل للشباب.	• دخل الأسرة المعيشية (بالدولار الأمريكي). • معدل بطالة الشباب المحلي (٪). • مستويات استهلاك الغذاء/ الأسرة من الطعام (بما في ذلك الخضار والفاكهة). • عدد ونسبة الشباب الذين يقدمون تقارير تفيد بأنهم متعلمين/مفتهم جيداً/تحوي منزلهم/البنية التحتية الأساسية على البنية التحتية الأساسية (المياه الجارية، الكهرباء، ... وغيرها). • عدد ونسبة الشباب الذين يقدمون تقارير تفيد بانخفاض معدلات النزاعات في السنة السابقة.
التوظيف المدعوم (مثل: الأتغال العامة وبرامج الخدمة العامة)	• نفس ما تقدم أعلاه	• عدد مشاريع البرامج الإزانية حسب النوع والموقع. • عدد البلديات التي تقدم الأتغال العامة/الخدمات.	• عدد المستفيدين الذين تم توظيفهم في كل نشاط. • عدد الوظائف المؤقتة التي تم إيجادها (حسب النوع والقطاع).	• عدد ونسبة الشباب الذين تحولوا إلى العمل الرسمي في غضون X شهرًا • عدد أيام العمل وساعات العمل الأسبوعية (حسب نوع النشاط). • متوسط الأجر بالساعة/باليوم/بالشهر.	• نفس ما تقدم أعلاه

جدول ٣-٣ أمثلة على المؤشرات الخاصة بمشاريع سبل عيش الشباب

نوع المشروع	المخرجات	الأنشطة	المخرجات	النتائج	نتائج المستوى العالي
خدمات التوظيف (مثال، دعم التنسيب الوظيفي)	نفس ما تقدم أعلاه	- عدد خدمات الإرشاد الوظيفي الذي تم تقديمها (في مكاتب العمل، في المدارس، ... وغيرها) - عدد جلسات الإرشاد الوظيفي التي تم تنظيمها - عدد معارض المهن والوظائف التي تم تنظيمها	- عدد الشباب المشاركين في خدمات إيجاد الوظيفة - عدد الشباب الذين تتطابق مؤهلاتهم مع احتياجات أرباب العمل - عدد الشركات والشباب المشاركين في معرض المهن/الوظائف المحلي	- عدد مقابلات العمل لكل مستفيد - عدد ونسبة الشباب الذين تم توظيفهم لمدة X شهر بعد التدخل - عدد ونسبة الشباب الذين تم الاحتفاظ بهم في العمل لمدة X شهر على الأقل	نفس ما تقدم أعلاه
مشاريع الشباب وتنظيم مشاريع الشباب	نفس ما تقدم أعلاه	- عدد مسابقات خطط العمل التي تم تنظيمها - عدد ساعات خدمات الدعم التي تم تقديمها - متوسط عدد ساعات المتابعة الأسبوع/الشهري	- عدد الشباب الذين قدموا خطط عمل كاملة - عدد المشاريع الشبابية التي يتم دعمها سنوياً - عدد ونسبة الشباب الذين يتحدثون إلى مستشارهم مرة واحدة كل أسبوعين على الأقل	- عدد ونسبة الشباب الذين بدأ مشروع شركة جديدة - عدد ونسبة الشركات المسجلة - إجمالي المبيعات خلال الأسبوع/الشهر الماضي - عدد الوظائف التي تم إيجادها - نسبة الأرباح التي تمت إعادة استثمارها	نفس ما تقدم أعلاه
الخدمات العالية المخصصة للشباب	نفس ما تقدم أعلاه	- عدد ورش العمل التي تم تنظيمها للمؤسسات المالية المشتركة - عدد مشاريع القروض الصغيرة التي تم إطلاقها للشباب - عدد حسابات التوفير الموجهة للشباب التي تم إنشاؤها	- عدد العاطلين الذين تلقوا التدريب في المؤسسات المالية الشريكة - عدد القروض التجارية الصادرة للشباب (حسب نوع المشروع) - متوسط حجم القروض - عدد حسابات التوفير التي تم فتحها للشباب	- معدل السداد السنوي - مبلغ المدخرات الحالية في (١) المصرف، (٢) مع مجموعات التوفير، (٣) في جميع المواقع الأخرى - عدد ونسبة الشباب الذين يضعون المدخرات جانباً بمجرد حصولهم على المال - عدد ونسبة الشباب الذين يفيدون بشعورهم بدرجة أفضل من الرضا عن وضعهم المالي	نفس ما تقدم أعلاه

الخطوة الثانية: جمع البيانات

لا يعتمد اختيار المؤشرات التي سيتم استخدامها في نظام المتابعة الخاص بنا على هيكل المشروع وأهدافه فقط، وإنما أيضاً على توفر البيانات وعلى الوقت والمهارات المطلوبة لجمع هذه البيانات. تشير البيانات إلى جميع أنواع المعلومات وليس فقط المعلومات القابلة للقياس الكمي.

اختر أسلوب جمع البيانات

يوجد أسلوبان واسعان لجمع البيانات: الأسلوب الكمي والأسلوب النوعي.

الأساليب الكمية تهدف إلى توفير صورة موضوعية لقياس حالة ما ببعض الطرق التي يتم تحديدها مسبقاً. توفر هذه الأساليب معلومات حول مجتمع الدراسة موضع الاهتمام في صيغة أبعاد مغلقة وكمية، بما في ذلك الصفات الديموغرافية والاقتصادية والاجتماعية وغير ذلك من الصفات. عادة ما تكون قائمة على أدوات منظمة ومعيارية تيسر التجميع والتحليل المقارن. تشمل الأمثلة الشائعة الاختبارات والمسوح والتعدادات. يتطلب إجراء الأساليب الكمية وجود مهارات في الإحصاءات.

الأساليب النوعية تهدف إلى فهم كيف ولماذا يفكر الناس ويتصرفون بالطريقة التي يتصرفون بها. تسعى الأساليب الكمية إلى فهم الأحداث من وجهات نظر أصحاب المصالح، وتسعى لتحليل كيف تفسر مجموعات الأفراد المختلفة تجاربهم وتبني على أساسها الواقع. تضم الأمثلة الشائعة على الأساليب الكمية المقابلات غير الهيكلية أو شبه الهيكلية ومجموعات النقاش البؤرية/المركزة والمراقبة المباشرة للمشاركين. يتطلب إجراء الأساليب النوعية تدريب في مجال الأنثروبولوجيا أو علم الاجتماع، وكذلك تدريب في إدارة أدوات تقييم معينة. تميل الأساليب النوعية لأن تكون أسرع في تنفيذها من الأساليب الكمية وعادة ما تكون أقل تكلفة.

تمتاز القواعد التي تحكم الإحصاءات بالشفافية ومن السهل نسبياً إتباعها، مما لا يتطلب حكماً مستقلاً من جانب المحلل. ونتيجة لذلك، عادة ما تحقق الأساليب الكمية مستويات أعلى من الثبات والصدق مقارنة بالأساليب النوعية. في المقابل، يعتبر تفسير البيانات النوعية مسألة إصدار حكم. ونتيجة لذلك، يكون من الأصعب تعميم الأساليب النوعية. ونظراً لمزايا كل من الفئتين والقيود المفروضة على كل منها، يوصى في كثير من الأحيان بالمرج بين الأساليب النوعية والكمية (نهج الأساليب المختلطة) للحصول على نظرة شاملة حول تنفيذ البرنامج وفاعليته. يقدم جدول ٣-٤ نظرة عامة على التقنيات الشائعة لجمع البيانات.

جدول ٣-٤ نظرة عامة على أساليب جمع البيانات

الأسلوب	الوصف	الإستخدام	المزايا	القيود
السجلات الإدارية وسجلات الإدارة	الوثائق التي تقدم معلومات حول عمليات إدارة المشروع.	لاختيار فاعلية إدارة المشروع أو تنفيذ الاستراتيجية.	تقدم معلومات حول العمليات التي يصعب تنفيذها من خلال وسائل أخرى.	- يكون البرنامج محددًا، وغير قابل لتعميمه. - تعتمد على مصداقية سجلات الإدارة.
الزيارات الميدانية (الجمع بين الملاحظة وإجراء المقابلات معا)	اختبار متعمق لموقع أو مكان محدد.	لمتابعة وفهم السياق.	- مستوى عال من التفصيل. - الوصول إلى بيانات الملاحظة.	- يكون البرنامج محددًا، وغير قابل لتعميمه. - تعتمد كثيراً على الوصول إلى المواقع الميدانية الملائمة.
مقابلات المبلغين الأساسيين (أصحاب المعرفة)	أسلوب يتم فيه الجمع المتعمق للبيانات مع أشخاص على قدر عال من الإطلاع.	للحصول على معلومات محددة ومفضلة كثيراً حول قضية معينة أو مجموعة من القضايا.	- مستو عال من التفاصيل. - يمكن أن تعالج مواضيع غير متوقعة. - تتميز بالمروية لاستكشاف القضايا بعمق. - يمكنها التقاط مجموعة من وجهات نظر أصحاب المصلحة.	- يكون البرنامج محددًا، وغير قابل لتعميمه. - تتغير الجودة كثيراً حسب مهارات الشخص الذي يجري المقابلة وعلى الراحة التي يشعر بها الشخص الذي تجري المقابلة معه.
مجموعات النقاش البورصة/المركزية	أسلوب لجمع المعلومات بشكل متعمق مع أعضاء مطلعين من فئة معينة من السكان (مثال، النساء، الشباب، كبار السن، العمال).	للحصول على معلومات محددة ومفضلة كثيراً حول وجهات نظر أصحاب المصالح بخصوص قضية معينة أو مجموعة من القضايا.	- كما في المقابلات التي تتم مع المبلغين الأساسيين. - تسمح بوجود تفاعل مع المشاركين وفيما بينهم.	- يكون البرنامج محددًا، وغير قابل لتعميمه. - تعتمد الجودة كثيراً على ديناميكية المجموعة (مثال، يمكن أن يتأثر المشاركون بالمشرف أو بأعضاء المجموعة المسيطر). - تحيات تتعلق بالتفسير. - التحليل الذي يستغرق وقتاً طويلاً.
الملاحظة المباشرة	أسلوب لجمع البيانات من خلال الملاحظة المباشرة (مثال، ملاحظة الفصل)، ويتم تسجيل المعلومات في سجل أو مفكرة.	للحصول على معلومات واقعية.	- مستوى عال من التفاصيل من ملاحظة محايد. - توفر معلومات حول السلوك الحقيقي بدلاً من السلوك الذي يتم التقرير عنه ذاتياً.	- غير قابلة لتعميمها. - هناك إمكانية كبيرة لانحياز القائم بالملاحظة. - تحيات تتعلق بالتفسير والترميز.

الأسلوب	الوصف	الاستخدام	المزايا	القيود
مراجعة السجلات الرسمية	وثائق رسمية توفر معلومات أساسية أو بيانات تاريخية حول ظواهر معينة	لاختيار العمليات الأساسية أو الاتجاهات/البيانات التاريخية لظواهر معينة.	- توفر معلومات يصعب الحصول عليها من خلال وسائل أخرى. - غير مكلفة.	- قيود محتملة في الوصول إليها. - يجب التحقق من صدق وثبات البيانات. - قد لا تكون البيانات ما هو مطلوب بالتحديد.
المسوح المصغرة	استبيان/ مسح مختصر يعمل على جمع مجموعة محدودة من البيانات.	للحصول على بيانات كمية حول عدد محدود من الأشخاص أو القضايا.	أسرع وأقل توسعاً من مسح الأسر	محدودة النطاق وبالتالي عادة ما لا يكون هذا الأسلوب مثقلاً.
مسوح الأسر	مجموعة واسعة من أسئلة المسح يمكن ترميز إجاباتها على الدوام.	للحصول على المعلومات حول عدد كبير من المستجيبين فيما يتعلق بوضعهم الاجتماعي والاقتصادي والبيانات الديموغرافية وأنماط الاستهلاك وغيرها.	- تقدم معلومات متعمقة حول مجتمع الدراسة موضع الاهتمام. - قابلة للتعميم أكثر من المسوح المصغرة. - قد تكون مصممة لجمع البيانات موضع اهتمام خاص.	- مكلفة. - تتطلب خبرة خاصة لتحقيق الصدق. - يصعب إقناع الناس بالاستجابة لاستبيانات طويلة.
مسوح لجان الخبراء	دراسة زمنية يتم فيها قياس المتغيرات على نفس الراحات مع مرور الوقت.	كما هو الحال بالنسبة لمسوح الأسر، مع اهتمام خاص بقياس التغيرات بمرور الوقت.	كما هو الحال بالنسبة لمسوح الأسر، يمكن أن تلتقط الديناميات على مدار فترة من الزمن.	كما هو الحال بالنسبة لمسوح الأسر، قد يكون هناك مشاكل تتعلق بالاحتفاظ بالمشاركين مع مرور الوقت.
التعداد	مسح لجميع السكان.	للحصول على مجموعة من المعلومات الكاملة حول مجتمع دراسة معين.	- قابلة لتعميمها - عادة ما تكون متوفرة من مصادر رسمية	- مكلفة - تستغرق وقتاً - غير متكررة أو مؤرخة

المصدر: مقتبس عن بيكر (٢٠٠٠): كريسويل (٢٠٠٨).

تكون آليات جمع البيانات أكثر أو أقل ملائمة للمستويات المختلفة في سلسلة النتائج. ستعتمد مؤشرات المدخلات ومؤشرات الإجراءات في المقام الأول على الإدارة وسجلات المشروع التي توضح استخدام المصادر وتنفيذ الأنشطة. يمكن للملاحظة المباشرة والزيارات الميدانية أن توفر البيانات لمؤشرات المخرجات، على سبيل المثال، عدد الشركات الصغيرة التي تم إنشاؤها. في كثير من الأحيان يتطلب قياس النتائج مزيجاً من المسوح الرسمية التي توفر المعلومات الكمية الموثوقة وكذلك الأساليب النوعية مثل المقابلات التي تجرى مع أشخاص مطلعين أو مجموعات النقاش البؤرية/المركزية لفهم الآليات الأساسية للتأثيرات إن حدثت وكيف تم تحقيقها. أخيراً، لأن نتائج المستوى الأعلى عادة ما تكون مرتبطة بتغيرات أوسع خارج السيطرة التامة للمشروع، يمكن للإحصاءات الرسمية أن تكون مفيدة عندما تكون متوفرة لمناطق جغرافية صغيرة (مثل البلديات) ويمكن أن تكون مفصلة حسب الخصائص الاجتماعية والديموغرافية.

تحديد تكرار وتوقيت جمع البيانات
ستعتمد الفترات الفاصلة بين أنشطة المتابعة على أهداف المتابعة. وكقاعدة عامة، كلما كان مستوى النتائج أرفع، كلما قلت حاجتنا إلى جمع البيانات، في حين يصبح عادة من الأصعب الحصول على معلومات دقيقة.

لتوضيح التكرار الأمثل لجمع البيانات، لنتخيل برنامجاً للتدريب الميداني يستمر لمدة ثلاثة أشهر. لإدارة التدريب بشكل فعال وكفؤ، نحتاج إلى معلومات حول عدد الموارد التي نستخدمها (من ميزانية ووقت العاملين والمواد وغيرها) وكيفية تطبيق أنشطتنا (عدد ساعات التدريب المقدمة كل أسبوع وعدد المشاركين وغيرها). قد نحتاج إلى جمع هذه المعلومات حول مدخلاتنا وأنشطتنا على نحو متكرر بعض الشيء، كل أسبوعين مثلاً.

أما تقييم مخرجاتنا (عدد المستفيدين الذين يتم تدريبهم بالفعل وخصائص المجموعة)، فسنحتاج إلى القيام به على نحو دوري، كل شهر مثلاً، وذلك بالرغم من أن هذه المعلومات ستعتمد على بيانات الحضور التي تم جمعها على مستوى يومي.

ولن يتضح ما إذا كان للتدريب أي أثر على النتائج (معرفة الشباب وقدرتهم على إيجاد الوظائف) إلا بعد انتهاء التدريب. يمكن قياس الآثار قصيرة الأجل، مثل زيادة المعرفة، في نهاية التدريب، فيما يجب أن تقاس الآثار التي تتطلب وقتاً أطول لتتجلى - مثل ما إذا كان قد تم تأمين وظائف للشباب - بعد ثلاثة إلى ستة أشهر بعد التدخل.

وأخيراً، من غير المحتمل أن تتجلى في العادة النتائج في المستوى الأعلى، مثل الزيادات في دخول الأسر أو الآثار الإيجابية الممتدة في أقل من عام (حسب سوق العمل المحلية) وبالتالي لا يمكن قياسها سوى على مدى فترات زمنية طويلة.

تحديد المسؤولين عن جمع البيانات
من المهم تحديد مسؤوليات جمع البيانات بوضوح. حيث سينتج عن عدم تحديد المسؤوليات على الأرجح فشل عملية جمع البيانات. عملياً، ستقع أنواع مختلفة من المتابعة تحت مسؤولية أفراد مختلفين، سواء ميدانياً وفي مقر الرئيسي للعمل. وفيما يلي الأشخاص المرجح إشراكهم في جمع البيانات:

- مدراء البرامج.
- أعضاء فريق المحلي في المشروع وموظفو المتابعة والتقييم.
- الشركاء المحليون في التطبيق (مثل المعلمين والمدربين وموظفي القروض).
- المستفيدون.
- أصحاب المصالح الآخرون (بما في ذلك الآباء وأفراد المجتمع).
- المتطوعون لجمع البيانات (مثل طلاب الجامعات).
- المستشارون من الخارج.
- شركات الدراسات والأبحاث.

[معلومة إرشادية]

استخدم الأساليب الكمية عند

- طلب بيانات رقمية أو بيانات قابلة لتعميمها لإقناع صناع القرار.
- تحتاج إلى معلومات ممثلة إحصائياً حول مجتمع الدراسة المستهدف ووضعهم وسلوكياتهم ومواقفهم.

استخدم الأساليب النوعية عند

- الحاجة لفهم أسئلة «كيف ولماذا»، أي عندما تكون هناك حاجة لشرح البيانات النوعية حسب الحافز والاتجاهات المؤثرة على السلوكيات.
- تفضيل النهج القائم على المشاركة.

[معلومة إرشادية]

يجب التخطيط لتوقيت جمع البيانات بالتوافق مع الواقع المحلي لكي لا تشكل عملية الجمع عبئاً على فرد أو أسرة ما. فلا يجب جمع البيانات في حين يخضع فيه الشباب إلى اختبارات مدرسية مثلاً أو عند الحاجة إلى الأيدي العاملة من الشباب خلال مواسم زراعية محددة.

[معلومة إرشادية]

ينبغي وضع تضارب المصالح في الاعتبار عند تحديد مسؤوليات جمع المعلومات والتقارير عنها. قد يكون لدى المعلمين والمدرسين على سبيل المثال حافز للغش فيما يتعلق بتسجيل المخرجات (مثل عدد ساعات التدريب المقدمة) أو النتائج (مثل عدد الشباب الذين تصنت درجات اختباراتهم أو وجدوا وظائف). لضمان صدق البيانات، نوصي (١) باللجوء إلى ملاحظين محايدين لضمان المتابعة المستقلة و(٢) التحقق من دقة المعلومات المقدمة، على الأقل بين الحين والآخر، من خلال زيارات مفاجئة للمواقع أو وسائل أخرى.

لمثال عن كيف حسنت المتابعة بالتصوير من حضور المعلمين وقللت من الحاجة إلى زيارات المتابعة في الهند، قم بزيارة <http://www.povertyactionlab.org/evaluation/encouraging-teacher-attendance-through-monitoring-cameras-rural-udaipur-india>

عند تحديد مسؤوليات جمع البيانات، حدد ماذا سيحدث للمعلومات عندما يتم جمعها. أضف إلى خطط جمع البيانات، تخزين البيانات وتصنيفها وتحليلها لضمان أن من يحتاج إلى المعلومات يمكنه الحصول عليها في أي وقت كان (أنظر نظام المتابعة ورفع التقارير أدناه).
لمعرفة المزيد عن المتابعة والتقييم القائم على المشاركة، راجع سايبو فلورز (٢٠٠٨)، باورز وتيفاني (٢٠٠٦) وغاولر (٢٠٠٥).

الخطوة ٣: صياغة المخاطر والافتراضات

ما هي العوامل الأساسية التي بإمكانها خفض الآثار المحتملة للمشروع وما هي الخطوات التي يمكن القيام بها للتخفيف من حدتها؟ في أي مشروع توجد عوامل لا يمكن التحكم بها وستؤثر على نجاح تدخلنا. ويمكن لهذه أن تتضمن عوامل مثل الطقس والاستقرار السياسي والحالة الأمنية المحلية والدعم من أصحاب المصالح المحليين. لذا، من المهم فهم هذه العوامل جيداً لتصميم المشروع وكذلك للمتابعة والتقييم.

تحديد الافتراضات خلال مرحلة التصميم
يمكننا تحديد الافتراضات عبر التفكير في العوامل التي تلعب دوراً هاماً في تحقيق أهدافنا في كل مرحلة من سلسلة النتائج وما الذي يمكن أن يؤثر على هذه العوامل (أنظر جدول ٣-٥). يمكن أحياناً أن تكون قد وردت المجموعة الأولى من الافتراضات في قسم المخاطر في خطة المشروع (Development Marketplace 2008).

جدول ٣-٥ أمثلة على افتراضات واستجابات المشروع

الافتراضات المحتملة	الخفة	تحت سيطرتنا؟ نعم/كلا
• يمكن إيجاد مدربين مستعدين للعمل في منطقة المشروع. • وكالة التوظيف مستعدة للشراكة.	المدخلات	• نعم، ولكن، لم يتم توظيفهم بعد. • أجل، تم توقيع مذكرة التفاهم بالفعل.
• تتوفر الكهرباء في موقع التدريب	الأنشطة	• كلا، ولكن، لم تقع مشاكل في الأشهر الأخيرة.
• يمكن للشباب حضور التدريب وليسوا مضطرين للعمل لإعالة عائلاتهم.	المخرجات	• كلا، ولكن، ستوزع قسائم للتعويض عن نقص الدخل.
• التدريب متصل باحتياجات سوق العمل ومقدم بجودة عالية.	النتائج	• أجل، لقد تم إجراء مسح لرأي أرباب العمل وسيتم متابعة أداء المدربين.
• يظل الاقتصاد المحلي (بما فيه الأسعار والرواتب في السوق) مستقرًا.	نتائج ذات مستوى عال	• كلا، لكن التوقعات إيجابية.

يجب إدراج الافتراضات التي لا يمكننا السيطرة عليها في مصفوفة النتائج في المستوى الذي تؤثر فيه. وبوجه عام، من المرجح أن تكون المدخلات والأنشطة تحت سيطرة المشروع أكثر من النتائج والنتائج ذات المستوى العالي.

متابعة الافتراضات أثناء تنفيذ المشروع لتوفير نظام تحذير مبكر حول القيود المحتملة والحلول الممكنة، يجب متابعة الافتراضات عن كثب. تسمح لنا متابعة الافتراضات بمعرفة كيف يمكن أن تؤثر في تطبيق المشروع ونتائجه، وبالتالي يمكنها مساعدتنا في تفسير انحرافنا عن أهدافنا وعلى اتخاذ التدابير التصحيحية.

تأسيس نظام المتابعة ورفع التقارير

التخطيط

بعد تطوير إطار عمل كامل ومنطقي يشتمل على مؤشرات وأدوات جمع بيانات وافتراضات، ستساعدك المهام التالية على الإعداد للمتابعة.

تصميم الأدوات الضرورية. ستسمح البيانات التي تم جمعها بشكل منهجي باستخدام أدوات مصممة جيداً بأن يتم صياغة التقارير بشكل سريع وموثوق، على أن يتم اختيار الأدوات على مجموعة من مجتمع الدراسة ذات صلة في خلال مرحلة التطوير ويجب أن ترشد نتائج هذا الاختبار تصميم الأدوات التابعة.

وضع إجراءات لحماية الشباب. على الرغم من أن الحكومات لا تشترط ذلك دائماً، غير أن المعايير المهنية تفرض أن تجري أنشطة جمع البيانات بطريقة تحمي حقوق ومصالح المشاركين. أما طبيعة هذه الإجراءات بالتحديد، فهي تخضع للمحددات المحلية، ولكن، يفضل القيام بالتالي كحد أدنى:

- إعداد الأدوات وإجراءات تدريب المقابلين بشكل يضمن عدم الكشف عن هوية المشاركين الشباب في البحث.
- الحصول على استمارات موافقة موقعة تتضمن تفاصيل المشروع والمخاطر المحتملة في حالة المشاركة. كما تفسر هذه الاستمارات بوضوح حقوق المشاركين، مثل حق التوقف عن المشاركة في عملية جمع البيانات متى رغبوا في ذلك. والحصول على الموافقة الشفوية ممن لا يجيدون القراءة.
- الحصول على موافقة من الآباء أو الأوصياء بالنسبة لمن لم هم دون السن القانونية التي تسمح لهم بالموافقة، ومن يعانون من إعاقة في النمو وغيرهم من المجموعات السكانية الأخرى المعرضة للخطر. وفي حالة عدم توفر الموافقة لأحد هؤلاء الأشخاص، يجب تجنب جمع البيانات عن الفرد المعرض للخطر.
- للحصول على مزيد من الإرشاد، أنظر قسم حماية مصالح المشاركين في بند ٧.
- جمع البيانات وفق الأساليب المختارة.** يجب الاستفادة قدر الإمكان من الإجراءات المتوفرة مثل تسجيل المشاركين أو سجلات الحضور للحد قدر الإمكان من عبء جمع البيانات على الموظفين والمشاركين.
- تطوير قاعدة البيانات.** إذا كانت البيانات التي تم جمعها معقدة، فقد يكون من المفيد تعيين شخص خبير في منصب مدير لقاعدة البيانات ليطور رموزاً وإجراءات من شأنها السماح لمستخدمين مختلفين بالحصول على المعلومات واستخلاص النتائج بقليل من التدريب. وتتوفر مجموعة متنوعة من نظم قواعد البيانات لذا يجب اختيار برنامج للحاسوب يوفر توازناً بين التحليل المتقدم وسهولة الاستخدام.

تجميع المعلومات وتحليلها

تعتمد أساليب تجميع المعلومات وتحليلها إلى حد كبير على الأساليب التي يتم استخدامها لمتابعة المشروع أو التدخل. لذلك، يجب البدء باتخاذ القرارات حول كيفية استخدام بيانات المتابعة في مرحلة مبكرة جداً من عملية تصميم المشروع. يجب أن يختار الفريق الأسلوب الأمثل لتنظيم وتحليل هذه البيانات بشكل فعال وكفؤ. ولتسهيل التحليل ورفع التقارير في المشاريع الكبيرة، يفضل إنشاء نظام للمعلومات الإدارية للربط بين كافة قواعد البيانات التي تستخدمها مختلف وحدات البرنامج.

[معلومة إرشادية]

تأكد من أن الأدوات المستخدمة تجمع أنواعاً مختلفة من المعلومات (عنوان السكن والبريد الإلكتروني ورقم الهاتف) من المستجيب وأيضاً من رفاقه وأفراد عائلته الذين بإمكانهم المساعدة في توجيهك إلى الشباب الذين يمتازون بسرعة التنقل. كما يمكن لقنوات التواصل الاجتماعي مثل فيسبوك أن تساعدك على التواصل مع الشباب ومتابعة أخبارهم.

[مصدر إلكتروني]

عينة من أدوات المسح، يتضمن بعضها استمارات الموافقة عنوان الموقع الإلكتروني <http://www.iyfnet.org/gpye-m&e-resource11>

[تعريف]

نظام إدارة المعلومات هو مزيج من تكنولوجيا الحاسوب والأشخاص والإجراءات يتم وضعها لجمع المعلومات وتنظيمها وتحليلها لدعم عملية اتخاذ القرارات. ويسمح بإدارة كميات ضخمة من البيانات بطريقة مركزية ومقارنة المؤشرات بحسب خصائص المستفيدين بمرور الوقت.

بالنسبة للبيانات النوعية، من الأفضل في كثير من الأحيان (وإن كان ذلك يشكل تحدياً من الناحية اللوجستية) استخدام برمجيات حاسوبية للتحليل النوعي. وهناك الكثير من الماركات للاختيار من بينها (مثل Atlas.ti، NVivo، أو MaxQDA) وكل منها يعمل بطرق مشابهة. وتسمح برمجيات التحليل النوعي للمستخدم باستيراد كافة المستندات ذات الصلة (مثل محاضر من مقابلات ومجموعات النقاش البؤرية/المركزة ومستندات المشروع والصور الفوتوغرافية)، ثم تطبيق مجموعة من الرموز المحددة مسبقاً. وبناء على درجة خبرة المستخدم، يمكن أن تستخدم الرموز كأداة تنظيم (عن طريق تجميع كافة الموضوعات المشابهة معاً من عدة مصادر) أو السماح بتحليل متطور يدرس العلاقات داخل هذه الموضوعات. ويتعين على الفريق اختيار البرنامج الحاسوبي الذي يلي احتياجاتهم من حيث خبرة العاملين والتكاليف.

بالنسبة للبيانات الكمية، عندما تسمح الموارد، يفضل في كثير من الأحيان استخدام عدد من النظم، ويجب أن يكون أحد هذه النظم قاعدة بيانات يوضح العلاقات بين المدخلات، مثل Microsoft Access. فقواعد البيانات تلك تسمح ببحث سهل وعرض للبيانات بالإضافة إلى عدد من المتغيرات المختلفة. وفي الواقع، فإن التحليلات التي تتم في قواعد البيانات من هذا النوع تكون وصفية بطبيعتها، حيث تقدم مقاييس النزعة المركزية (مثل المتوسط والوسيط والمنوال والانحراف المعياري). وإذا تطلب المشروع والأدوات كانت مصممة ومدارة بطريقة من شأنها أن تسمح بتحليل أكثر تطوراً، فقد يرغب العاملون في المتابعة في استخدام حزمة برنامج حاسوبي إحصائي مثل SPSS أو Stata. وبالإضافة إلى حزم البرمجيات الحاسوبية الإحصائية المتاحة عموماً والتي تعتمد على محرك القرص الصلب لحاسب آلي واحد، هناك أيضاً استخدام متزايد لبرنامج «Cloud»، القائم على نظم إدارة البيانات والتحليل، والذي يسمح لفريق ضخم بالتعاون في مهام المتابعة والتحليل.

التعلم وصنع القرار

لن يكون للمتابعة سوى قيمة ضئيلة إذا لم نتعلم من ونعمل بناء على البيانات الناتجة عن التحليل. فكون الشخص في حلقة مستمرة من العمل والتفكير، فإن ذلك يساعد على تذكر أن الأوضاع تتغير، وأن احتياجات المستفيدين من المشروع قد تتغير، وأن الاستراتيجيات وأنشطة المشروع تحتاج إلى إعادة النظر فيها وتعديلها. فالمؤسسات والمشروعات تصبح خامدة عندما لا تتعلم والمتابعة الصارمة تجبرنا على الاستمرار في التعلم شابيرو (٢٠٠٣).

وطبقاً لـ شابيرو (٢٠٠٣)، تنطوي ترجمة التعلم إلى إجراء على ما يلي:

- النظر إلى التبعات المحتملة للتحليل على برنامجنا.
- سرد قائمة بخيارات العمل.
- مناقشة الخيارات مع أصحاب المصالح من الداخل والخارج والتوصل إلى اتفاق مشترك والحصول على تكليف بالمهمة لاتخاذ إجراء.
- مشاركة التعديلات والخطط مع بقية المنظمة، وعند الضرورة، مع مانحين والمستفيدين من برنامجنا.
- تنفيذ الخطوة.
- متابعة الآثار.

في عام ٢٠١١، بدأت منظمة شباب الأعمال الدولية (YBI)، وهي شبكة تضم أكثر من أربعة وثلاثين برنامجاً مستقلاً لتنظيم المشروعات الشبابية حول العالم، تنفيذ نظام إدارة العمليات (OMS) العالمية المستند لبرنامج cloud لأغراض المتابعة. ويسمح نظام إدارة العمليات لأعضاء منظمة شباب الأعمال الدولية بتتبع وتحليل نطاق عريض من مؤشرات الأداء الرئيسية المتصلة بالكفاءة التنظيمية والنتائج. ويمكن تقييم جودة ملف قروض العضو ونجاح الأعمال التجارية الخاصة بمنظمي الأعمال التابعين له، مقابل عوامل مثل الخصائص الاجتماعية- الديموغرافية لمنظم الأعمال، والإرشاد والتدريب الوظيفي الذي حصل عليه، وشروط القرض. ويساعد هذا النظام على زيادة الدقة وتيسير التجميع الآني للمعلومات من قبل الفريق المركزي لشبكة منظمة شباب الأعمال الدولية.

رفع التقارير

بصفة عامة، كلما ارتفع مستوى الجماهير التي نستهدفها في التسلسل الهرمي لمنظمة ما، كلما قل ما نحتاج أن نقدمه من تفاصيل وتفسيرات كثيرة عند نقل النتائج التي توصلنا إليها. فتقديم رسائل واضحة مدعومة ببيانات مجمعة ومعلومات مختصرة تميل أن تكون أكثر ملائمة للجماهير رفيعة المستوى ممن يهتمون في الغالب بالصورة الأشمل. ويمكننا أن نصيغ تقاريرنا بشكل يناسب كل جمهور (انظر جدول ٣-٦).

جدول ٣-٦ تصميم تقارير خاصة للجمهور

الجمهور المستهدف	الصيغة	التوقيت/التكرار
العاملون بالمشروع	عرض شفهي وإحصائيات مكتوبة موجزة في اجتماعات الفريق.	أسبوعياً
فريق الإدارة	تقارير مكتوبة وعرض شفهي.	شهرياً
الشركاء	عرض شفهي وإحصائيات مكتوبة موجزة.	شهرياً
المانحون	يعتمد ذلك على متطلبات المانح. عادة ما تكون تقارير موجزة مكتوبة تبرز تقدم المشروع، والمشكلات التي تم مواجهتها والنتائج والأثر وفاعلية التدخل/الاستراتيجية وغيرها.	ربع سنوي/نصف سنوي

يجب دوماً رفع تقارير بيانات المتابعة مقارنة بنقطة أساس جمع هذه البيانات والقيم المستهدفة وتقديمها في شكل «بسيط، وواضح وقابل للفهم بسهولة» (كيوسك وريست ٢٠٠٤، ص. ١٣٣). يمكن أن تكون الأدوات بصرية مثل الرسوم البيانية والمخططات والخرائط مفيدة لإبراز البيانات والرسائل الرئيسية.

الموارد

قد تكون أنظمة المتابعة باهظة التكلفة. فبالإضافة إلى التكاليف الثابتة (أجهزة وبرامج الحاسوب والعاملين)، فهناك أيضاً تكاليف متغيرة والتي تشمل تدريب موظفي التعداد المحليين، والتعاقد مع مستشارين من الخارج، ونشر النتائج (انظر جدول ٣-٧). فمن المهم أن يتم تخصيص الأموال بشكل ملائم لنظام متابعة وتقييم مشروع ما طبقاً لميزانية وأخذها في الاعتبار في أية خطة استراتيجية. وغالباً ما يكون الحال عندما يتم التحقق من التكاليف، أن يتردد مدراء المشروع في إنفاق موارد ضخمة على نظام المتابعة والتقييم، والذي يبدو أنه يكون على حساب أنشطة التدخل. ومع ذلك، فبدون نظم متابعة مناسبة، يجري تنفيذ البرنامج مع احتمال ضعف أدائه أو فشله، مع وعي ضئيل بهذه المشاكل. كذلك، بدون متابعة، لا يمكننا اقتناص الفرص التي يتم تحقيق نجاحات كبيرة فيها. وفي النهاية، تعتبر نظم المتابعة هامة للغاية لإدارة المشروع ومكون حاسم في أي تدخل.

التكاليف الثابتة	
تكلفة العاملين	- المقر الرئيسي: نسبة الوقت التي يقضيها منسق المتابعة والتقييم في إدارة نظام المتابعة والتقييم، والذي يتراوح بين ١٠ بالمائة إلى ١٠٠ بالمائة، بناءً على حجم المشروع. - محلياً: عادةً من ٥٠ - ١٠٠ بالمائة من وقت المسؤول المحلي لنظام المتابعة والتقييم في إدارة تنفيذ أنشطة المتابعة والتقييم، بالإضافة إلى صغار موظفي الدعم.
الأجهزة والمعدات	الحواسيب الآلية ومسجلات الصوت والكاميرات وغيرها.
البرمجيات	تراخيص أدوات التحليل الكمي والنوعي.
التكاليف المتغيرة	
التدريب	بناء قدرات العاملين وموظفي التعداد وأعضاء المجتمع وغيرهم.
السفر والانتقالات	سفر موظفي المقر الرئيسي إلى الميدان لإجراء معاينات دورية وتقديم المساعدة الفنية، بالإضافة إلى سفر الموظفين المحليين إلى المواقع الميدانية لضمان التنفيذ المعياري لأنشطة المتابعة والتقييم.
جمع البيانات وتحليلها	التعاقد مع بائعين طرف ثالث مثل شركات الدراسات.
المستشارين	التعاقد مع الخبراء من الخارج للقيام بمهام محددة.
الطباعة	المستندات والتقارير، وغيرها.

النقاط الرئيسية

١. يجب أن يكون لكل تدخل نظام متابعة راسخ ليتمكن من التتبع المستمر للتنفيذ والنتائج، بغض النظر عما إذا ما كان سيتم تقييم المشروع من عدمه.
٢. يتعين على مدراء المشروع وأصحاب المصالح الرئيسيين مجتمعين أن يطوروا سلسلة نتائج لتحديد أسباب التدخل بشكل واضح والتعرف على المؤشرات الرئيسية وآليات جمع البيانات والافتراضات.
٣. يقدم نظام المتابعة معلومات مستمرة عن اتجاه التغيير وسرعته وحجمه. وهو أيضاً يتيح لنا التعرف على التطورات غير المتوقعة في المشروع أو بيئته. وذلك يوفر الأساس لمعرفة ما إذا كان التدخل يسير في الاتجاه المقصود منه ويجعل المتابعة الجيدة هامة للإدارة الفاعلة للمشروع.
٤. بيانات المتابعة بيانات وصفية ولا توضح بالضرورة سبب وكيفية حدوث تغييرات محددة. كما أنها أيضاً لا توفر أساساً لإسناد التغييرات الملحوظة إلى التدخل. بمعنى آخر أنها لا تثبت حدوث هذه التغييرات بسبب برنامجنا.

دراسة حالة NUSAF: نظام المتابعة



بغرض بناء أساس لتفسير نتائج تقييم الأثر، كان من المهم بالنسبة لبرنامج NUSAF أن يكون لديه معلومات جيدة حول إذا ما كان برنامج فرص الشباب كان قد تم تنفيذه بالشكل المنشود. وبناء عليه، استخدم برنامج NUSAF مزيج من الأدوات الكمية والنوعية لتتبع الأنشطة والمخرجات. على سبيل المثال، عندما تم توزيع المنح النقدية على المجموعات الشبابية من خلال الحكومة المركزية، تم سؤال الشباب إذا ما كانوا قد حصلوا بالفعل على تلك الأموال من عدمه. ثم تم مقارنة هذه المعلومات مع السجلات الحكومية.

كذلك حاول البرنامج فهم توزيع الأموال واستخدامها داخل المجموعة. وحيث كان الهدف من صرف الأموال هو التدريب وشراء المواد والأدوات، تتبع برنامج NUSAF نسب الحضور، وعدد وقيمة الأدوات والمواد الخاصة بهم، سواء بدأوا عمل تجاري أم لا، وإن ما زالوا يديرون العمل التجاري أم لا.

بالرغم من أن هذه المعلومات لم توفر إجابات تتعلق بأثر البرنامج، إلا أنها ساعدت مسؤولي البرنامج وموظفي المتابعة والمقيمين على فهم ما إذا تم تنفيذ البرنامج طبقاً لما هو مخطط له من عدمه وكيفية تأثيره على المشاركين فيه. وساعد هذا الفهم أيضاً أثناء تحليل نتائج التقييم، مثلاً على توضيح سبب استفادة بعض المشاركين من البرنامج بدرجة مختلفة عن الآخرين.

المصدر: باتلمان ومارتينيز وفيالا (٢٠٠٩).

قراءات رئيسية

لجنة المانحين لتنمية المشروعات. ٢٠١٠.

The DCED Standard for Measuring Achievements in Private Sector

Development, Control Points and Compliance Criteria. الإصدار الخامس.

<http://www.enterprise-development.org/page/measuring-and-reporting-results>

كيوسك، جيه. زد، وريست، آر. سي. ٢٠٠٤.

Ten Steps to a Results Based Monitoring and Evaluation System:

A Handbook for Development Practitioners. واشنطن، دي سي: البنك الدولي. أنظر

الفصول ٢-٦. <http://www.oecd.org/dataoecd/23/27/35281194.pdf>

بند ٤: اختيار النوع الصحيح من التقييم

أخطر الأخطاء لا يتم ارتكابها نتيجة إجابات خاطئة.
وإنما الشيء الخطير حقاً هو طرح السؤال الخطأ.

— بيتر دراكر

على الرغم من أهمية وجود نظام متابعة جيد لمعرفة ما إذا كان تدخلنا يتحرك في الاتجاه المقصود أم لا، إلا أنه لا يجب بالضرورة عن السؤال: كيف أو لماذا تنتج التغييرات، ولا يثبت أن تدخلنا هو السبب في أية تغييرات ملحوظة في النتائج. ونحن نحتاج إلى إجراء تقييمات من أجل استكمال المعلومات التي حصلنا عليها من نظام المتابعة الخاص بنا. فالتقييمات عبارة عن تقديرات دورية لأهمية تدخلنا وكفاءته وفاعليته وتأثيره واستدامته. وسوف يعتمد نوع التقييم الأنسب لمشروعنا أولاً على احتياجاتنا من المعلومات. وبالتالي، تتمثل الخطوة الأولى في أي تقييم في تحديد ما نريد أن نتعلمه. وهذه الأهداف التعليمية بالإضافة إلى السياق التشغيلي الخاص بنا، بدوره، سوف يحدد أي نوع من التقييم يصلح لبرنامجنا.

ما هو الغرض من التقييم؟

كخطوة أولى لاتخاذ قرار حول ما إذا كان التقييم ضرورياً وأي تصميم يجب اختياره، فمن المهم أن نحدد بوضوح ما الذي نريد الحصول عليه من التقييم. ما هو القرار الذي سوف يتم التوصل إليه عن طريق التقييم وما هي أنواع المعلومات اللازمة لاتخاذ هذا القرار؟ هل نرغب في أن نعرف أكثر إلى أي مدى يتم تنفيذ برامجنا بشكل جيد، وما إذا كانت برامجنا تحقق أهدافها أم لا، وما إذا كان المستفيدون منا في وضع أفضل بالفعل نتيجة لتدخلنا؟ وبصفتنا مدراء البرنامج ومقيمي، يجب أن نحدد أولاً أسئلتنا وأهداف التعلم الخاصة بنا، ثم نختار أداة التقييم الأكثر ملائمة لتوفير المعلومات الضرورية (كارلان ٢٠٠٩).

وبشكل عام، تتناول التقييمات ثلاثة أنواع من الأسئلة (إيماس وريست ٢٠٠٩):

- أسئلة وصفية تحاول وصف العمليات والشروط والعلاقات التنظيمية ورؤى أصحاب المصالح (ما الذي يحدث في مشروعنا؟).
- أسئلة معيارية تقارن الذي يحدث بما يجب أن يحدث. حيث تقارن الوضع الحالي بالأهداف والغايات التي سبق تحديدها (هل تم تنفيذ مشروعنا على النحو المنشود؟ هل يتم إنجازها كما هو متوقع؟).
- أسئلة الأسباب والآثار تفحص النتائج وتحاول أن تقيس الفرق الذي أحدثه التدخل. فهي تحاول التوصل إلى ما إذا كانت الأهداف قد تحققت كنتيجة لمشروعنا (ما هو الأثر أو العلاقة السببية لبرنامجنا على النتائج ذات الأهمية؟).

وفي النهاية، فالأمر متروك لنا بالنسبة لما يجب أن نطرحه من أي من الأسئلة السابقة، بناء على التدخل المحدد.

- ترتيب الأسئلة التي نطرحها. من الناحية العملية، قد يكون لدينا العديد من الأسئلة حول مختلف الفئات التي نود أن نجيب عنها. تعد سلسلة النتائج التي وضعناها (أنظر جدول ٤-١) أحد السبل الفاعلة لترتيب جميع أسئلة التقييم الممكنة. وفي واقع الأمر، إذا قمنا بوضع نظام متابعة جيد (أنظر بند ٣)، فيجب أن يكون هناك اتفاق حول منطق تنفيذ المشروع ونتائجه، مما ييسر من مهمة تحديد أهداف التعلم الهامة في جميع مراحل التدخل. يمكن أن ترتبط الأسئلة الوصفية والمعارية بجميع مراحل سلسلة النتائج، في حين تشير أسئلة الأسباب والآثار خاصة إلى النتائج والنتائج ذات مستوى عال.

وصفية	المدخلات	الأنشطة	المخرجات	النتائج	نتائج ذات مستوى عال
• ما وجه المقارنة بين تكلفة البرنامج والتدخلات المماثلة؟ • ما هي المؤهلات الواجب توافرها لدى مزودي الخدمة؟ • ما هي التدخلات الأخرى القائمة؟	• هل يعلم الشباب بالبرنامج وكيف يتم تأهيلهم للالتحاق به؟ • ما هي آليات تقديم الخدمة المستخدمة حالياً؟ • إلى أي مدى قد يختلف تطبيق البرنامج من موقع لآخر؟	• كم عدد الشباب المشاركين (حسب العمر والجنس ... وغيره)؟ • من الذي ينسحب من المشاركة؟ • ما هي أكثر الخدمات المستخدمة؟	• هل يشعر المشاركون بالرضا تجاه البرنامج؟ • هل طرأ على أي من المشاركين أية تغيرات ملحوظة فيما يخص معلوماتهم أو اتجاهاتهم... وما إلى ذلك؟ • كم عدد المشاركين بالبرنامج الذين يجدون وظيفة خلال ٣ أشهر؟	• هل ترتفع معدلات البطالة بين الشباب على المستوى المحلي أم تنخفض؟ • هل يزداد مستوى دخل الأسرة؟	
• هل تسير معدلات الإنفاق لدينا وفق ما حددناه في الميزانية؟ • هل يفي العاملون والموارد المالية بالحاجة؟ • هل يكرر البرنامج جهود أخرى؟	• هل تتم عملية اختيار المشاركين بشكل عادل ومنصف؟ • هل تأخر تطبيق البرنامج؟ • هل يتم إتباع الدلائل التنفيذية؟	• هل نحقق المساواة المنشودة بين الجنسين؟ • هل سنتمكن من تحقيق هدف تدريب ٥٠٠٠ شاب سنوياً؟	• هل يتزايد دخل الشباب بنسبة ٢٠٪ وفقاً للخطة؟ • هل تمكن ٨٠٪ من المستفيدين من التدريب الحصول على وظائف خلال ثلاثة أشهر، وفقاً للمطلوب إنجازه؟ • ما هي التأثيرات السلبية أو الإيجابية التي حدثت دون تخطيط مسبق لها، إن وجدت؟	• هل استطاعت المزيد من الأسر تحقيق الاكتفاء الذاتي؟ • هل استطاعت المزيد من الأسر تحقيق الأمن الغذائي؟	
• لا يوجد	• لا يوجد	• لا يوجد	• هل تمكن المشاركون نتيجة حصولهم على تدريب التوظيف من الحصول على وظائف ذات معدلات دخل تفوق التي كان من المتوقع الحصول عليها إذا لم يتلقوا هذا التدريب؟ • هل تزيد البرامج التدريبية من فاعلية التدريب الفني المقدم؟ • هل يختلف تأثير البرنامج على الأولاد عن تأثيره على البنات؟	• هل يسهم المشروع في الحد من الفقر في المنطقة؟ • ما هي الآثار الأخرى الناجمة عن هذا التدخل على مستويات المعيشة في المجتمع الأوسع نطاقاً؟	

العلاقة بين معايير التقييم وأسئلة التقييم. توجد طريقة أخرى للتفكير في أسئلة التقييم وذلك من خلال التفكير في المعايير العامة للتقييم والتي حددتها في بادئ الأمر منظمة التعاون الاقتصادي والتنمية (OECD). وكما ذكرنا، فإننا نعني بالتقييمات إجراء تقديرات بصفة دورية لكل من أهمية وكفاءة وفاعلية وآثر واستدامة تدخلنا (منظمة التعاون الاقتصادي والتنمية ١٩٩١). إذا أمعنا النظر، نجد أن الأهمية والكفاءة والفاعلية ترتبط في المقام الأول بالأسئلة المعيارية، في حين يرتبط الأثر بأسئلة الأسباب والآثار. أما الأسئلة المرتبطة بالاستدامة فتكون إما أسئلة معيارية (هل قد يستمر التدخل بعد انتهاء التمويل من قبل المانحين؟) أو أسئلة أسباب وآثار (هل تعد الآثار المرصودة مستدامة على مر الزمن؟)، ولا تعد أيًا من هذه الأسئلة وصفية محضة، غير أن الأسئلة المعيارية قد تنطوي بشكل طبيعي على أسئلة وصفية. ينظم جدول ٢-٤ كل معيار من معايير التقييم وما يناسبه من أسئلة التقييم.

جدول ٤-٢ الصلة بين معايير التقييم وأسئلة التقييم

المعيار	الوصف	التفاصيل	نوع أسئلة التقييم
الأهمية	هل تتناسب الأهداف مع المشكلات أو الاحتياجات التي يتم تناولها؟	• إلى أي مدى لا تزال أهداف البرنامج صادقة؟ • هل تتفق أنشطة ومخرجات البرنامج مع الشكل الإجمالي لتحقيق هذه الأهداف؟	معيارية
الكفاءة	هل تم تسليم المشروع في موعده المقرر وبطريقة ناجحة ومجدية اقتصادياً؟	• هل ينفذ البرنامج أو المشروع باستخدام أكثر الطرق كفاءة؟ • ما هي التكلفة التي نتحملها لكل مخرج أو مستفيد، وكيف نستطيع مقارنة هذه التكاليف بالتدخلات المماثلة؟	معيارية
الفاعلية	إلى أي مدى حقق التدخل أهدافه؟	• إلى أي مدى استطعنا تحقيق النتائج المنشودة؟ • ما هي أهم العوامل المؤثرة في تحقيق الأهداف أو عدم تحقيقها؟	معيارية
الأثر	ما هي التغيرات الإيجابية والسلبية الناتجة جراء التدخل؟	• ما هي النتائج ذات المستوى العالي الناتجة عن البرنامج أو المشروع؟ • ما هو الفرق الحقيقي الذي أحدثه النشاط لدى المستفيدين؟	أسباب وآثار
الاستدامة	هل هناك فوائد مستمرة بعد انتهاء التدخل؟	• إلى أي مدى قد تستمر الفوائد الناتجة عن المشروع بعد توقف تمويل المانحين؟ • ما هي أهم العوامل التي أثرت في تحقيق الاستدامة أو عدم تحقيقها؟	معيارية أو أسباب وآثار

المصدر: استناداً إلى منظمة التعاون الاقتصادي والتنمية (بدون تاريخ).

تحديد أولويات ما نطرحه من أسئلة. لا يمكن لأي نوع من الأسئلة أن يحل محل غيره من الأنواع، إلا أن الأسئلة المعيارية عادة ما تشتمل على الأسئلة الوصفية وتقوم عليها. فكل نوع من هذه الأسئلة يعطي جانب من جوانب المشروع المختلفة ويقدم نوع مختلف من المعلومات والتي قد تكون مفيدة. فإذا أرادنا التركيز على النتائج، سنحتاج عندها لأسئلة الأسباب والآثار على وجه الخصوص. ففي الواقع، إذا كانت أهدافنا تتمثل في تحديد التدخلات المتعلقة بسبل عيش الشباب الواعد وإثبات التأثيرات الناجمة بالفعل عن تدخلنا، فستحتل أسئلة الأسباب والآثار جزءاً من أهداف التعلم الخاصة ببرامجنا، إن لم تكن لها الأولوية.

يؤدي كل نوع من أنواع الأسئلة الثلاث - الوصفية والمعيارية والأسباب والآثار - إلى اعتبارات مختلفة فيما يتعلق بنوع التقييم المقرر وضعه. يقوم مدراء البرنامج ومقيميهم بتخصيص سؤال محتمل لكل نوع من أنواع الأسئلة الثلاث، ثم يقومون بالنظر في تبعات كل نوع من الأسئلة لوضع تصميم التقييم. وبهذا، فإننا عندما نختار مجموعة من أسئلة التقييم، فإننا نحدد قائمة أدوات المتابعة والتقييم المناسبة التي من خلالها يتسنى لنا الإجابة هذه الأسئلة (معايير الرقابة الحكومية (GAO) ١٩٩١).

ربط أسئلة التقييم بتصميم التقييم

لا يوجد نموذج تقييم «يناسب الجميع»، فخلاصة القول، أن اختيار التقييم يجب أن يعتمد بشكل أساسي على الأسئلة السابقة، وليس على تفضيلاتنا المنهجية أو تفضيلات المقيمين الداخليين أو الخارجييين. رغم أن هذا الأمر قد يبدو واضحاً، إلا أنه ليس شائعاً ومتبعاً في معظم الأحوال.

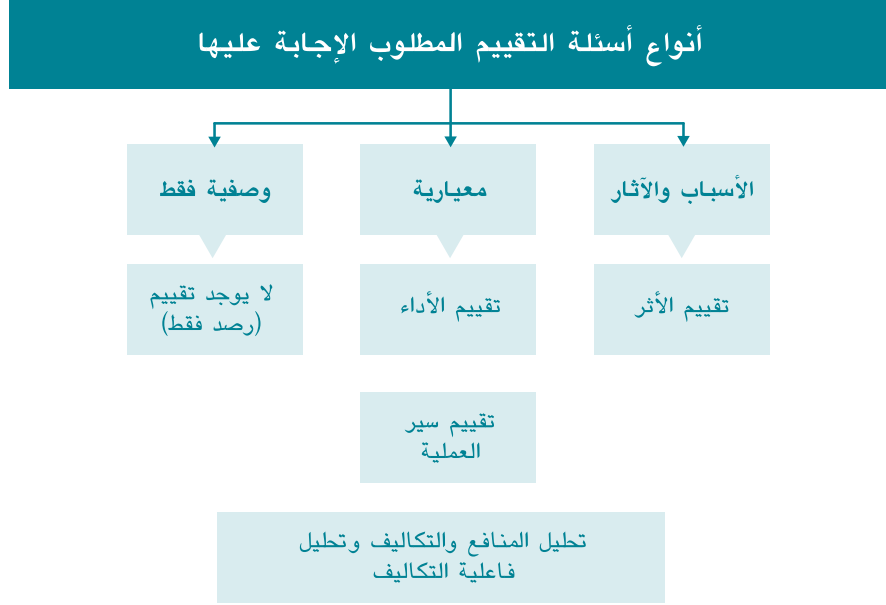
[معلومة إرشادية]

تأكد من تحديد الجمهور المراد تقييمه وما الذي يريد هذا الجمهور معرفته. قد يقوم بعض العاملين بالبرنامج أو مدراءه بطلب بعض التقييمات داخل المنظمة، بينما قد يطلب المانحون أو صانعو السياسات تقييمات أخرى. قد تختلف احتياجات المعلومات الداخلية والخارجية، بما يؤدي إلى اختلاف أسئلة التقييم. وبالتالي فمن المهم إشراك أصحاب المصلحة في تحديد أسئلة التقييم والأسئلة ذات الأولوية.

المصدر: مقتبس من Rubio (٢٠١١).

يقدم شكل ١-٤ نظرة عامة على خيارات التقييم بحسب نوع الأسئلة الذي نود منحه الأولوية.^٢

شكل ١-٤ من أسئلة التقييم إلى تصميم التقييم



لا يوجد تقييم

إذا طلب مدير البرنامج الحصول على معلومات وصفية عن التدخل، على سبيل المثال حيث يكون المشروع في مرحلة البداية الأولى، وكان الهدف يتمثل في الحصول على بعض المعلومات العامة عن كيفية تنفيذ البرنامج، فلن يكون التقييم الكامل ضرورياً. وفي هذه الحالة، قد تكون المعلومات التي يتم الحصول عليها من المتابعة كافية. ومن الواضح أنه لبلوغ هذه الغاية يتطلب توافر نظام متابعة فعال ويعمل جيداً مبني وفقاً لسلسلة نتائج محددة بوضوح، بالإضافة إلى المؤشرات وأدوات جمع البيانات وغيرها من هذا القبيل (أنظر بند ٣). إذا ما قمنا بتطبيق مثل هذا النظام، ستتوافر لدينا معلومات وصفية عن البرنامج بطريقة سهلة نسبياً.

تقييم الأداء

تقوم تقييمات الأداء بتقدير مدى جودة صياغة أهداف البرنامج (أنظر المعايير في بند ٢)، بالإضافة إلى تقييم مدى تقدم البرنامج في هذه تحقيق هذه الأهداف (رويو ٢٠١١). كما أنها تنظر أيضاً في مدى ملائمة إطار النتائج القائم، وما إذا كانت هناك تناقضات بين الموارد والأنشطة والأهداف، وما إذا كان ينبغي تعديل الأولويات أو الجداول الزمنية من أجل تحقيق الأهداف المنشودة بشكل أفضل. يمكن القيام بمثل هذه التقييمات في كافة مراحل التطبيق، إلا أنها تعد مألوفة وشائعة في مراجعات منتصف المدة (عندما ينصب تركيزها على التعلم لصالح إدارة البرنامج) وكذلك في نهاية البرنامج (حيث تركز على المسألة والدروس المستفادة لإجراء التدخل في المستقبل). يمكن إجراء تقييمات الأداء، والتي عادة ما ينفذها مقيم مستقل، بشكل سريع نسبياً وبتكلفة متواضعة، حيث تقوم بشكل كبير على إجراء البحوث المكتبية والمقابلات المختارة. على الرغم من ذلك، قد تنطوي تقييمات الأداء في بعض الأحيان على عمليات جمع معلومات مستفيضة، مثل مقارنة نتائج المشاركين قبل وبعد المشاركة أو أدوات نوعية أخرى. وعلى الرغم من أهمية تقييمات الأداء لأغراض تقييم الجودة بصفة عامة، إلا أنها لا توفر ثقة مطلقة حول ما إذا كانت التغييرات الملحوظة قد حدثت بسبب تدخل معين.

٢ ١ هناك أنواع أخرى من التقييم تركز على مستويات أخرى من تقديم المساعدة (بما في ذلك القطاعات والموضوعات وفاعلية المساعدة) والتي لم نشير إليها في هذا البند. يقتصر هذا البند على تقييم المشروعات والبرامج.

تقييم سير العملية

على خلاف تقييمات الأداء التي تركز بشكل أساسي على تحقيق الأهداف، تسعى تقييمات سير العملية إلى فهم كيفية عمل البرنامج فهماً كاملاً وتُسعى لتقدير مدى جودة تطبيق البرنامج. حيث تحدد ما إذا كان هناك فجوات بين الأنشطة والمخرجات المقررة والمحققة، وتحاول فهم أسباب الفجوات بينها. استناداً إلى المعلومات الوصفية مثل ما هي الأنشطة المقدمة ومن هم الأشخاص المشاركون (أو من الذي لا يشارك)، فإنها تحدد وسائل تحسين جودة الخدمات المقدمة. تجرى تقييمات سير العملية على مراحل محددة لتمثل بذلك نظام إنذار مبكر، أو قد تتم هذه التقييمات في حالة متابعة مشكلات كتأخر التنفيذ أو عدم شعور المستفيدين بالرضا وذلك من خلال إجراءات المتابعة القياسية (البنك الدولي ٢٠٠٢). تعتمد تقييمات سير العملية إلى الاعتماد على مزيج من الأدوات النوعية والكمية، بما في ذلك مقابلات المبلغين الرئيسيين، وإجراء استبيان لمعرفة مدى رضا المشاركين والملاحظات المباشرة ومجموعات النقاش البؤرية/المركزة.

تقييم الأثر

تعد تقييمات الأثر النوع الوحيد من التقييمات للإجابة تحديداً عن أسئلة الأسباب والآثار بطريقة قابلة للقياس. وتتطلب منا مثل هذه الأسئلة ليس فقط تحديد ما إذا كانت النتائج المرغوبة بها قد تحققت أم لا بل وكذلك ما إذا كانت هذه النتائج قد تحققت بفضل تنفيذ البرنامج أم لا. وكما أشار غيرتير وزملاؤه (٢٠١١، ص. ٤)، يعد هذا التركيز على السببية والإسناد «السمة المميزة لتقييمات الأثر» ويحدد مجموعة الأساليب التي يمكن استخدامها. (يقدم بند ٦ نظرة عامة على الأدوات المناسبة). ومن أجل تقدير التأثير السببي للبرنامج على النتائج موضع الاهتمام، يجب أن يقوم أي أسلوب يتم اختياره على تقدير ما يسمى بالواقع المضاد، بمعنى ماذا كان سيحدث للمشاركين في البرنامج في ظل عدم وجوده. وللقيام بذلك، تتطلب تقييمات الأثر إيجاد مجموعة مقارنة؛ وهي تتكون من مجموعة من الأفراد كان سيكون لهم نتائج مماثلة لتلقي البرنامج، في ظل عدم وجود التدخلات (دفلو وجليزيستر وكريمير ٢٠٠٦). وهذا ما يجعل تقييمات الأثر مختلفة عن غيرها من التقييمات الأخرى. وبالنتيجة، تتطلب تقييمات الأثر وقتاً أكبر ومهارات كمية كما أنها عادة ما تكون أكثر تكلفة من غيرها من أنواع التقييم الأخرى. واستناداً إلى المعلومات التي تقدمها، فهي مهمة في المساعدة على تحديد إجابات لأسئلة استراتيجية، كتوسيع نطاق التدخلات الفاعلة أو استبعاد البرامج غير الواعدة (روبيو ٢٠١١). علاوة على ذلك، تزيد هذه التقييمات من قاعدة المعرفة العالمية بشأن الفاعلية النسبية لمختلف أنواع تدخلات سبل العيش للوصول إلى نتائج محددة ومساعدتنا في فهم أية خيارات تصميم البرامج (الجرعة وأسلوب تقديم الخدمة وغيرها) هي الأكثر أهمية ضمن فئة برنامج معينة.

[تعريف]

الواقع المضاد يشير إلى النتائج التقديرية للمشاركين في البرنامج وذلك في غياب البرنامج. ويجيب الواقع المضاد عن سؤال ماذا كان يمكن أن يحدث للمستفيد في حالة عدم وجود البرنامج.

تحليل فعالية التكاليف وتحليل المنافع والتكاليف

تعمل تقييمات فعالية التكاليف وتقييمات المنافع والتكاليف على تقدير تكاليف البرنامج النقدية وغير النقدية ومقارنتها باستخدامات بديلة لنفس الموارد والمنافع الناتجة عن التدخل (بيكر ٢٠٠٠). ويقاس تحليل فعالية التكاليف (CEA) التكلفة حسب المخرج أو النتيجة (مثلاً: ٣٠٠ دولار لكل شاب يتم تدريبه و ٥٠٠ دولار لكل وظيفة يتم إيجادها) ويقارن هذه التكلفة بتدخلات مشابهة خاصة بنا والمنظمات الأخرى. وبهذا، يجيب التحليل عن السؤال حول المخرج أو الناتج الذي نحصل عليه مقابل كل دولار يتم إنفاقه (وصفي) وما إذا كان يوجد فجوة مع توقعاتنا (معياري). وبالمقابل، يقارن تحليل العائد والتكاليف (CBA) بين إجمالي التكاليف المتوقعة وإجمالي المنافع (النتائج) المتوقعة لتدخل ما، حيث يتم عادة التعبير عن التكاليف والمنافع بلغة النقد. فعلى سبيل المثال، إذا كان يفترض برنامجنا مساعدة ٥٠٠ شاب على إيجاد وظائف والحفاظ عليها أو إقامة مشاريع صغيرة مستدامة، فإننا سنعمل على (١) تقدير المنافع الإجمالية من حيث دخول أعلى وصحة أفضل وجريمة أقل وغيرها، و(٢) مقارنة هذه المنافع مع التكاليف الإجمالية للتدخل. وبما أن تحليل المنافع والتكاليف ينظر لقيمة الفوائد التي تم تحقيقها، فإنه يتطلب تقديراً موثقاً لدرجة تأثير البرنامج على النتائج موضع الاهتمام، مما يجعله مفيداً للغاية بالاشتراك مع تقييمات الأثر (من أجل الحصول على وصف أكثر تفصيلاً، أنظر بند ٨). يعطي مربع ٤-١ روابط لأمثلة على أنواع التقييم التي تم مناقشتها أعلاه.

تقييمات الأداء

- Human Sciences Research Council. 2007. *Mid-term Review of the Expanded Public Works Programme: Synthesis Report*. Pretoria: Southern Africa Labour and Development Research Unit, University of Cape Town; Rutgers School of Law; and ITT (UK).
http://www.hsrc.ac.za/research/output/outputDocuments/5465_Hemson_MidtermreviewofEPWPsynthesisreport.pdf
- Education and Employment Alliance. 2010. *An Evaluation of Partnerships in Support of Youth Employability: Global Report*. <http://www.iyfnet.org/document/1436>

تقييمات سير العملية

- Miller, E., and MacGillivray, L. 2002. *Youth Offender Demonstration Project Process Evaluation*. Chapel Hill: Research and Evaluation Associates Inc.
http://wdr.doleta.gov/opr/fulltext/YODP_final.pdf
- The Lewin Group, Inc. 2003. *Evaluation Design for the Ticket to Work Program—Preliminary Process Evaluation*.
<http://www.lewin.com/content/publications/2526.pdf>

تقييمات الأثر

- Attanasio, O., Kugler, A. and Meghir, C. 2009. "Subsidizing Vocational Training for Disadvantaged Youth in Developing Countries: Evidence from a Randomized Trial." IZA Discussion Paper No. 4251. Bonn: IZA. <http://ssrn.com/abstract=1426738>
- Mensch, B., Grant, M., Sebastian, M., Hewett, P., and Huntington, D. 2004. "The Effect of a Livelihoods Intervention in an Urban Slum in India: Do Vocational Counseling and Training Alter the Attitudes and Behavior of Adolescent Girls?" Working Paper No. 124, New York: The Population Council.
<http://www.popcouncil.org/pdfs/wp/194.pdf>

تحاليل فعالية التكاليف وتحاليل المنافع والتكاليف

- Elias, V., Nunez, F., Cossa, R., and Bravo, D. 2004. *An Econometric Cost-Benefit Analysis of Argentina's Youth Training Program*. Washington, DC: IDB.
<http://www.iadb.org/res/publications/pubfiles/pubR-482.pdf>
- Jastrzab, J., Masker, J., Blomquist, J., and Orr, L. 1996. *Evaluation of National and Community Service Programs—Impacts of Service: Final Report on the Evaluation of American Conservation and Youth Service Corps*. Bethesda, MD: Abt Associates Inc.
http://www.abtassociates.com/reports/ccc_youth_0596.pdf
(Note: This is an impact evaluation and a cost-benefit analysis combined.)

هل يلاءم سياقنا التشغيلي نوع التقييم المنشود؟

وكما أشارت (معايير الرقابة الحكومية (GAO) ١٩٩١، ص. ١٥)، «إن الاتفاق على الأسئلة التي لها الأولوية الأعلى واختيار تصميم التقييم هو أمر، بينما معرفة ما إذا كانت الأسئلة يمكن الإجابة عنها، وإن كانت كذلك، ما هي التكلفة من حيث المال والعاملين والوقت هو أمر آخر تماماً.» بعد صياغة الأسئلة المناسبة وتحديد النوع المحتمل للتقييم، نحتاج إلى تقييم السياق التشغيلي للتدخل لفهم ما هو التقييم الذي يمكن تنفيذه ضمن القيود المحددة.

التوقيت

ترتبط الأسئلة المتعلقة بنوع المعلومات المطلوبة بشكل وثيق بالسؤال حول متى يجب أن تتوفر نتائج التقييم. حيث تحدد معرفة الوقت الذي يجب أن تتوفر فيه هذه النتائج الوقت الذي يجب جمع المعلومات فيه.

متى يتم تحديد طلب إجراء تقييم؟

يعطي التخطيط المبكر مرونة أكثر فيما يتعلق باختيار أداة تقييم ملائمة. فعلى سبيل المثال، تحتاج العديد من أساليب تقييم الأثر لأن يتم التخطيط لها حتى قبل بدء التنفيذ. ويفضل أن يكون التخطيط للتقييم جزءاً من تخطيط البرنامج («تقييم مستقبلي»). ولكن في كثير من الحالات، قد تنشأ الحاجة إلى معلومات فجأة، نتيجة - على سبيل المثال - لمشاكل مفاجئة على أرض الواقع، أو طلب من مانح. وبالمثل، يمكن للقيود التشغيلية، مثل التنفيذ السريع من أجل إنفاق الأموال، أن تملئ الجدول الزمني للتقييم. وعلى الرغم من أن هذه القيود لا مفر منها في الحياة الواقعية، إلا أنها تقلل من الخيارات المتاحة للتقييم التي يمكن أن تكون متاحة في ظل هذه الظروف.

في أية مرحلة من مراحل البرنامج تظهر الحاجة إلى المعلومات؟

تختلف الاحتياجات للمعلومات حسب دورة حياة البرنامج. فعلى سبيل المثال، قد يتطلب البرنامج الذي تم تخطيطه للتو تحليلاً لفعالية التكاليف للمساعدة في تحديد ما إذا كان سيتم تنفيذ البرنامج أم لا. ومن ناحية أخرى، بالنسبة لتدخل أطلق حديثاً، قد نحتاج إلى معرفة مدى إتباع إجراءات البرنامج وما إذا كان يلزم القيام بأية تعديلات ضرورية لضمان تشغيل ناجح للمشروع مستقبلاً (روبيو ٢٠١١). وفي كثير من الأحيان، يمكن تقدير هذه الاحتياجات للمعلومات حتى قبل بدء البرنامج، وبالمثل يمكن تقدير التوقيت التقريبي للتقييم.

كم يستغرق التقييم؟

تعتمد المدة التي يستغرقها التقييم جزئياً على الأساليب المتبعة لجمع وتحليل البيانات، والتي تختلف حسب نوع التقييم، كما تعتمد على الاتساع والعمق المطلوبين لدراسة معينة وهما أمران يختلفان ضمن كل نوع من أنواع التقييم. وبشكل عام، فإنه من الإنصاف افتراض أنه يمكن القيام بتقييمات الأداء الصرف خلال شهر إلى ثلاثة أشهر، بما أنها تعتمد اعتماداً كبيراً على البحوث المكتبية وعدد محدود من المقابلات. وبالمقابل فإن تقييمات سير العملية يمكن أن تختلف اختلافاً كبيراً في النطاق، فقد تكون بنفس سرعة تقييمات الأداء، ولكن يمكن أن تستغرق حتى ستة أشهر أو أكثر عندما يتم تحليل عمليات معقدة. أما بالنسبة لتقييمات الأثر فتكون على الأرجح الأكثر استنفاداً للوقت (سنة أشهر وحتى سنتين أو أكثر)، بما أن منهجيتها تحتاج إلى أن تكون مخططة بشكل جيد وقد يوجد حاجة لجمع بيانات جديدة. ويمكن أن يستغرق تحليل المنافع والتكاليف نفسه أقل من شهر إذا توفرت جميع البيانات الضرورية. أما إذا كانت هناك حاجة لجمع المعلومات أولاً، فيمكن أن يستغرق وقتاً أطول بكثير.

يوضح مربع ٤-٢ عند أية نقطة في دورة حياة البرنامج تكون أفضل لإجراء استراتيجيات التقييم المختلفة.



المصدر: مقتبس من البنك الدولي (٢٠٠٧ إصدار الربع الثاني، ص. ١٨).

المرحلة ١. على وشك أن تبدأ التجربة الأولى لتدخل مبتكر وغير مختبر نسبياً متعلق بسبل عيش الشباب. ما هو التقييم الذي يجب استخدامه؟

في المراحل المبكرة من برنامج ما، نحتاج عادة إلى التأكد أن كل شيء يتم القيام به حسبما هو مخطط. ولا ينصح بإجراء تقييم أثر في هذا الوقت لأن النتائج لن تعكس الجودة الحقيقية للبرنامج. فمن المناسب أكثر التركيز على المتابعة وتقييم سير العملية حتى تتحقق القدرة التشغيلية الكاملة للبرنامج ويتم حل مشاكل التنفيذ الشائعة في تأسيس برامج جديدة. ويمكن أن تكون أساليب جمع البيانات النوعية (مثل المقابلات التي تتم مع المبلّغين الرئيسيين ومجموعات النقاش البؤرية/المركزة) مفيدة على وجه التحديد في هذه المراحل المبكرة حيث إنها يمكن أن تجيب عن سؤال لماذا تعمل بعض العناصر أو لا تعمل حسب الغاية المنشودة. ويتم عادة الإشارة لهذه المرحلة التجريبية الأولية للبرنامج بـ «دراسة الجدوى» من أجل الحصول على «إثبات المفاهيم»، أي للتحقق مما إذا كان بالإمكان تنفيذ البرنامج فعلياً حسب ما هو مخطط له.

المرحلة ٢. التدخل قائم منذ سنة، وتم حل المشاكل التشغيلية المبكرة. وأظهرت المتابعة أن المستفيدين راضون عن البرنامج. فهل يجب علينا توسيع البرنامج أو تكراره في مكان آخر؟

قد يكون حان الوقت الآن لتقييم للأثر. البرنامج بدأ ويعمل ونحن واثقون من جودة التنفيذ. سيسمح لنا تقييم الأثر بالتأكد على أن البرنامج أثر على النتائج موضع الاهتمام. كما يمكننا أيضاً استخدام تقييم الأثر لمقارنة فاعلية بدائل تصميم البرنامج (مثل: توليفات مختلفة من الأنشطة وكثافات مختلفة للأنشطة) إذا كنا لا نزال غير واثقين من عناصر معينة في التصميم. كما سيساعدنا التقييم في فهم بعض الآثار المحتملة غير المرغوبة (الإيجابية أو السلبية). ونتيجة للمعلومات التي يتم الحصول عليها من تقييم الأثر، يمكننا اتخاذ القرار حول ما إذا كان ينبغي استثمار أموال وافرة في البرنامج أم لا.

المرحلة ٣. أسفر تقييم الأثر عن نتائج كلية إيجابية للغاية. فهل لا زلنا نحتاج لإجراء تقييم؟

على الرغم من أن النتائج الإيجابية لا تقتضي ضمناً أن البرنامج سيعمل بنفس الطريقة الجيدة في سياقات مختلفة، لكن يمكننا الآن أن نكون واثقين إلى حد ما من دقة نظرية التغيير ومجموعة الأنشطة الخاصة بنا. وهذا أساس جيد لتوسيع البرنامج بحيث يضم عدداً أكبر من المشاركين أو تكراره في مواقع مشابهة. ولن يكون من الضروري على الأغلب إجراء تقييم آخر للأثر إلا إذا رغبنا في إدخال تعديل هام على تدخلنا. ومع ذلك، يجب أن نكون متأكدين أن تظل جودة التنفيذ مرتفعة وأننا نحقق أهدافنا. ويجب أن تبقى المتابعة على جميع المستويات، بما فيها متابعة النتائج، مكوناً أساسياً في برنامجنا. فضلاً عن ذلك، يمكن أن تساعد تقييمات الأداء المستقلة التي تتم على فترات منتظمة في التحقق من الملائمة المستمرة للبرنامج وجودته.

الموارد

قد لا يكون بالإمكان تنفيذ بعض أساليب التقييم المرغوبة الأخرى إذا لم نكن نمتلك الموارد البشرية والمالية المطلوبة لتنفيذها. فمن المهم تقييم المهارات والتمويل المتوفرين في برنامجنا أو منظمنا للتأكد من توافقهما مع احتياجات التقييم الذي نتوقعه.

المهارات

يتطلب إجراء تقييمات الجودة مهارات خاصة قد لا تكون موجودة دائماً في برنامج ما أو منظمة ما. وفي هذه الحالة ومن أجل ضمان الحيادية، يكون غالباً من المفيد استخدام مقيمين خارجيين. يلخص جدول ٣-٤ بعض أهم المهارات المطلوبة لإجراء الأنواع المختلفة من التقييمات.

جدول ٣-٤ المهارات المطلوبة طبقاً لنوع التقييم

المهارة	الوصف	الأداء	سير العملية	الأثر	تحليل المنفعة والتكلفة
تصميم البرنامج والمتابعة	- الاطلاع في مجال وضع برامج سبل عيش الشباب. - الخبرة في تصميم البرامج. - معرفة عامة بأساليب جمع البيانات الكمية والنوعية. - معرفة البلد. - شهادة جامعية في العلوم الاجتماعية.	!	!	✓	لا يوجد
جمع البيانات الكمية	- تدريب متخصص في تصميم المسوح وتطبيقها ميدانياً. - بعض المعرفة في تحليل البيانات الكمية. - مهارات في إدارة البرامج لتشكيل فريق من موظفي التعداد وقيادته. - شهادة جامعية في العلوم الاجتماعية.	✓	✓	!	لا يوجد
تحليل البيانات الكمية	- تدريب متخصص في الإحصاء والاقتصاد القياسي. - درجة الماجستير أو الدكتوراه في الاقتصاد أو الصحة العامة أو في مجال ذي صلة.	✓	✓	!	✓
جمع البيانات النوعية	- تدريب متخصص في تطبيق الأساليب النوعية. - شهادة الماجستير أو الدكتوراه في علم الاجتماع أو الأنثروبولوجي أو علم النفس.	✓	!	✓	لا يوجد
تحليل البيانات النوعية	- تدريب متخصص في ترميز البيانات النوعية وتحليلها. - شهادة الماجستير أو الدكتوراه في علم الاجتماع أو الأنثروبولوجي أو علم النفس.	✓	!	✓	لا يوجد
التقدير	- تدريب متخصص في تقدير تكاليف ومنافع برامج الخدمات الإنسانية. - شهادة الماجستير أو الدكتوراه في الاقتصاد أو الصحة العامة أو في مجال ذي صلة.	✓	لا يوجد	✓	!

! مطلوب؛ ✓ يفضل

التمويل

تولد الاختلافات في نطاق وأشكال جمع البيانات المتنوعة وتحليلها مجموعة واسعة من تكاليف التقييم. عادة ما يكون الاستناد إلى البحوث المكتبية والمقابلات مع المبلغين الرئيسيين أرخص بكثير من تصميم المسوح الجديدة وإدارتها مع عدد كبير من الأفراد. لذلك فإن تقييمات الأداء هي بالتالي عادة أرخص أنواع التقييم، في حين تعد تقييمات الأثر الأكثر كلفة (أنظر جدول ٤-٤).

جدول ٤-٤: تقديرات التكاليف لأنواع التقييم المختلفة

نوع التقييم	التكلفة	العوامل التي تؤثر على التكلفة
تقييم الأداء	١٠,٠٠٠ - ٣٠,٠٠٠ دولار	نطاق التقييم وراتب المقيم
تقييم سير العملية	١٠,٠٠٠ - ٦٠,٠٠٠ دولار	كما في تقييم الأداء، لكنه غالباً ما يستخدم المزيد من أدوات جمع البيانات، لذا، يمكن للتقييم أن يستغرق وقتاً أطول.
تقييم الأثر	١٥,٠٠٠ - أكثر من مليون دولار	تختلف التكلفة كثيراً من برنامج إلى آخر حسب الأسلوب المستخدم: كلما زادت البيانات المجمعة، ازدادت تكلفة التقييم (أنظر البندين ٦ و٧ لمزيد من التفاصيل).
تحليل المنافع والتكاليف وتحليل فعالية التكاليف	١٠,٠٠٠ - ٣٠,٠٠٠ دولار	يستند إلى ما إذا كان قد سبق أن تم قياس المنافع وما إذا كانت البيانات جاهزة للاستخدام

عندما تكون كل البيانات جاهزة بالفعل، يمكن ألا تتجاوز تكلفة تقييمات الأثر ١٥,٠٠٠ دولار، وذلك على الرغم من أن تكلفتها تتجاوز في معظم الحالات ١٠٠,٠٠٠ دولار. قد تبدو تقييمات الأثر غير عملية للبرامج ذات الميزانيات المتواضعة، غير أنه يمكن تبرير تكلفتها إذا كان التدخل ينفذ أو سينفذ لفترة طويلة وعلى نطاق واسع. علاوة على ذلك، ليس من الضروري أن تتحمل المنظمة التي تطبق البرنامج دائماً تكلفة تقييم الأثر كاملة، بل يمكنها تقديم طلب للمساعدة المالية لتقوم بالتقييمات (أنظر بند ٧ لمزيد من التفاصيل حول وضع ميزانية تقييم الأثر).

السياق السياسي

قد يكون لأصحاب المصالح المختلفين داخل المنظمة وخارجها مصالح متناقضة فيما إذا ما كان يتوجب إجراء تقييم ما أو لا والمسائل التي يجب دراستها ونوع التقييم وأسلوبه واستراتيجية جمع البيانات ومن يجب استخدامه للقيام بالتقييم إذا تم الإقرار بالقيام به. قد تؤدي كل هذه العوامل إلى ضغوطات على اختيار التقييم وتؤثر في أهمية البحث المخطط وجودته. وقد تتراوح مثل هذه الضغوطات ما بين تلميحات إلى أنه لا يجب دراسة مسائل معينة وعدم موافقة السلطات الرسمية رسمياً على مقابلة بعض مجموعات الشباب أو العائلات أو المجتمعات.

لذا، فمن المهم محاولة فهم المصالح المختلفة والبيئة السياسية القائمة في سياق محدد. وتساعد الأسئلة التالية في البدء بالتحليل:

- ما هو السياق السياسي المحلي وتوزيع السلطات؟
- ما هي العلاقات التي تربط المستفيدين ومدراء المشاريع وصناع السياسات والمناحين وغيرهم من أصحاب المصالح؟
- ما هي المصالح والحوافز التي تواجه كل مجموعة من أصحاب المصالح في التأثير على سير التقييم وتصميم البرنامج؟ إذا كان البرنامج يستهدف مثلاً مجموعة معينة من الشباب على نحو وثيق، سيكون لأولئك الذين لم يتم إشراكهم حافزاً للتأثير في البرنامج وفي تقييمه بطريقة تمكنهم أيضاً من الحصول على الفوائد.
- إذا أظهر التقييم أثراً ما، من هم الراحون والخاسرون المحتملون من أي إصلاح برامجي أو سياسي قد يتأتى عن التقييم؟
- هل ستسمح البيئة المحلية بإجراء تقييم صارم ومستقل، وهل ستدعم المقيمين لينشروا نتائجهم المستندة إلى الأدلة بغض النظر عن العواقب السياسية؟

قررت منظمة دولية غير حكومية أن تجري مع شريكها المحلية في البرازيل تقييماً لأثر برنامج تدريب على تحسين فرص التوظيف كانتا تطبقانه معاً. بعد كثير من المشاورات، تم الاتفاق على متطلبات الأهلية، بما فيه اختيار المشاركين بشكل عشوائي. غير أن الشريك المحلي كان سبق أن عقد اتفاقاً مع شركة خاصة أرادت أن تؤثر في القرارات بشأن أي من الشباب سيشاركون في البرنامج، ما سيتسبب في انحياز نتائج أي تقييم. هذا الصراع جعل من غير الممكن إجراء الدراسة بشكل فعال.

قد يساعدنا العمل على فهم ما يقلق أصحاب المصالح من خلال التفاعل المستمر والمفتوح في تحديد طرق التعامل مع الضغوطات والمصالح المتضاربة وفي حشد الدعم للتقييم. كما أنه من المفيد عادة إحضار مقيمين من الخارج، فبالإضافة إلى أنهم سيحضرون معهم مجموعة من المهارات المحددة، سيكون من الأسهل المحافظة على استقلالية تقييمهم.

أنواع البرامج التي تبرر عادة إجراء تقييم للأثر

مع أنه يمكن لتقييمات الأداء وسير العملية وتحليل فعالية التكاليف أن تشكل جزءاً من كل برنامج، إلا أنه يجب إجراء تقييمات الأثر وتحليل المنافع والتكاليف على نحو انتقائي أكثر. فوفقاً لـ غيرتير وزملائه (٢٠١١)، يتم تعبئة الجهود الإضافية والموارد اللازمة لإجراء تقييمات الأثر على نحو أفضل عندما يكون البرنامج (١) ذا أهمية ومؤثر استراتيجياً، (٢) ومبتكراً أو غير مختبر، (٣) وقابلاً لتكرار تطبيقه.

ذو أهمية ومؤثر استراتيجياً

كم ستكون النتائج مهمة ومفيدة للبرامج والسياسات والحوارات المرتبطة بالسياسات المستقبلية؟ إذا كانت مجازفات التدخل عالية - لأن البرنامج يتطلب مثلاً موارد وتغطيات ضخمة، أو يمكن التوسع ليضم عدداً كبيراً من الناس - فيجب عندئذ التفكير في إجراء تقييم للأثر. وينطبق ذلك على المبادرات الجديدة كما على البرامج القائمة عندما نحتاج إلى اتخاذ القرارات حول إكمالها أو توسيع نطاقها أو إنهاؤها. في الواقع، حتى إذا كان تقييم الأثر باهظ الثمن، إلا أنه قد يكون فعال من حيث التكلفة بما أن نتائجه ستساعد في التوصل إلى تحسينات هامة في أداء البرنامج. وفي الحقيقة، في حالة المبادرات الكبيرة، حتى التحسينات الصغيرة قد ينتج عنها توفيرات كبيرة للمنظمة التي تطبقها (البنك الدولي ٢٠٠٩).

مبتكر أو غير مجرب

ما هي الحالة الراهنة للأدلة أو المعرفة حول الآثار المقترحة للبرنامج؟ إذا كنا نعلم القليل عن فاعلية نوع التدخل، سواء على المستوى الدولي أو في سياق محدد، يمكن لتقييم الأثر أن يضيف معرفة كبيرة إلى منظمتنا وإلى الحقل بأكمله. وهذا شأن معظم برامج سبل عيش الشباب التي ما زالت قاعدة أدلتها صغيرة. وفي حالة عدم توفر أية أدلة أو كونها ضئيلة جداً، عادة ما يوصى بالبداية ببرنامج اختبري يضم تقييماً للأثر. وحتى لو توفرت الأدلة حول نوع معين من التدخل، قد لا يزال تقييم الأثر مبرراً إذا كان البرنامج سيطبق في سياق مختلف أو إذا كان سيتضمن جوانب مبتكرة لم يتم اختبارها من قبل.

[مصدر إلكتروني]

الفجوات المعرفية وأسئلة البحث
الممكنة لتقييم الأثر
عنوان الموقع الإلكتروني
<http://www.iyfnet.org/gpye-m&e-resource3>

في حين يجب تفسير التعميمات التالية بحذر، إلا أننا نرى أنه يبدو أن الأدلة المتوفرة حول برامج سبل عيش الشباب تكون ضعيفة في هذه المجالات بالتحديد:

أنواع البرامج: معظم التقييمات متوفرة في مجال التدريب وتنمية المهارات، فيما لا تزال الأدلة حول أنواع التدخلات الأخرى كلها مثل توظيف الشباب المدعوم وخدمات التوظيف وتنظيم المشاريع الشبابية والخدمات المالية المقتصرة على الشباب والبرامج التي تستهدف المجموعات المستبعدة تكون قليلة نسبياً.

مميزات التصميم: قليل ما يعرف عن الفاعلية النسبية لبدائل البرامج. في كل نوع من البرامج، ما هو أثر إتباع مكونات مختلفة للبرنامج أو أساليب تربوية وجرعات وأساليب تقديم الخدمات مختلفة؟

السياق: إن الأدلة حول برامج سبل عيش الشباب قليلة بشكل خاص في الشرق الأوسط وشمال أفريقيا وآسيا وجنوب الصحراء الإفريقية. إضافة إلى ذلك، ما زلنا بحاجة إلى مزيد من الأدلة فيما يخص التدخلات وميزات التصميم التي تلاءم كل من السياقين الريفي والحضري أكثر والإطارين الرسمي وغير الرسمي أو بينات الدول الخارجة من صراع أو الضعيفة.

المستفيدون: كيف تؤثر أنواع مختلفة من البرامج في الشباب بشكل مختلف بحسب فئة عمرهم ونوعهم ومستوى تعليمهم وخلفيتهم الاجتماعية الاقتصادية؟ وما الذي ينجح مع المجموعات المحرومة؟ وما هي التداعيات الإيجابية أو السلبية للتدخلات المتعلقة بسبل العيش على الأقران والعائلات والمجتمعات؟

النتائج: ما هي آثار برامج سبل العيش ليس على التوظيف وسوق العمل فحسب، بل على التصرفات الخطيرة أيضاً والمشاركة المدنية وتكون العائلة والصحة العقلية وغيرها؟ تكاد الأدلة على الآثار طويلة الأمد لمعظم التدخلات تكون غير متواجدة أيضاً.

لمراجعة الأدلة المتوفرة، قم بزيارة سجل وظائف الشباب (www.youth-employment-inventory.org) وكنغهام وسانشين-بيورتا وويرملي (٢٠١٠).

قابل لتكرار التطبيق

إلى أي مدى وفي ظل أية ظروف يمكن لبرنامج اختبائي ناجح أو صغير النطاق أن يتم توسيع نطاقه أو يكرر تطبيقه على مجموعات مستهدفة مختلفة؟ إذا كان تصميم التدخل محدد للغاية ويستهدف سياقاً ضيقاً ومحددًا، يكفي على الأرجح إجراء تقييم لسير العملية من شأنه أن يساهم في تيسير التطبيق. ولكن، إذا كان من الممكن تعديل حجم البرنامج أو تطبيقه في إطار مختلف، يصبح تقييم الأثر عندئذ خطوة مهمة في توفيره لتبرير كاف لإعادة تطبيق البرنامج.

يقدم جدول ٤-٥ جدولاً يلخص أنواع التقييم.

جدول ٤-٥ نظرة عامة على أنواع التقييم الرئيسية

تحليلات فعالية التكاليف والمنافع والتكاليف	تقييم الأثر	تقييم سير العملية	تقييم الأداء	ما هي الأسئلة الرئيسية التي يمكن لهذا النوع من التقييم الإجابة عنها؟
- هل تكاليف البرنامج مبررة مقارنة بالتدخلات المشابهة؟ - هل إجمالي تكاليف البرنامج مبررة من حيث ما تم تحقيقه من فوائد؟	- كيف تغيرت رفاحية المشاركين نتيجة للتدخل؟ - هل هناك أية عواقب غير مقصودة، إيجابية أو سلبية، على المشاركين في البرنامج؟	- هل الموارد والنظم المناسبة (الإدارة والمعلومات وغيرها) في موضعها الصحيح؟ - هل يتم تنفيذ البرنامج طبقاً للتصميم؟ - ما هي الخطوات الفعلية والأنشطة الواجب أخذها لتقديم منتج أو خدمة؟ - ما الذي يعرفه المستفيدون أو غيرهم من أصحاب المصالح عن البرنامج؟	- هل للبرامج أهداف واضحة؟ - هل تصميم البرنامج ملائم لتحقيق الأهداف؟ - إلى أي مدى تحققت أهداف البرنامج؟ - هل يلزم تغيير الأولويات؟	ما هي الأسئلة الرئيسية التي يمكن لهذا النوع من التقييم الإجابة عنها؟
يتم إجراء التقييم بصفة عامة أثناء تخطيط مسبق لتحديد ما إذا كان البرنامج يستحق التنفيذ أو الاستمرار فيه أم لا أو بعد الانتهاء من البرنامج لتحديد التكاليف النهائية.	يجب تصميم التقييم أثناء التخطيط للبرنامج، ولكن لن تكون النتائج النهائية متاحة عادة إلا بعد الانتهاء من (مرحلة) البرنامج.	يمكن إجراؤه في أي وقت، مرة واحدة أو بشكل دوري، للتأكد أن تنفيذ البرنامج يجري في مساره الصحيح أو لفهم مخاوف أو مشاكل تشغيلية محددة.	يمكن إجراؤه في المراحل المبكرة من تنفيذ البرنامج، أو لمراجعة متوسطة الأجل، أو عند الانتهاء من البرنامج.	متى يمكن إجراء هذا التقييم؟
من ١ إلى ٣ شهور	- سنة شهور على الأقل (تقييم بأثر رجعي) - من ١٢ إلى ٢٤ شهر (تقييم متوقع)	من ١ إلى ٦ شهور	من ١ إلى ٣ شهور (أو أكثر إذا كان يتم استخدام تخطيط التقييم قبلي / بعدي)	كم يستغرق من الوقت؟
مراجعة مكتبية لوثائق البرنامج المتاحة والمؤلفات ذات الصلة بالإضافة إلى مقابلات المبلغين الأساسيين.	تحليل إحصائي واقتصاد قياسي لمسوح وبيانات إدارية، جنباً إلى جنب مع تحليل البيانات النوعية.	مزيغ من المقابلات مع العاملين بالبرنامج والعلماء ومسوح آراء المستخدمين ومراجعة السجلات والملاحظة المباشرة ومجموعات التركيز وتحليل بيانات المتابعة.	مراجعة مكتبية للوثائق القائمة وزيارات ميدانية مختارة، يتم استكمالها ربما عن طريق متابعة تحليل البيانات ومقابلات مع المستفيدين وأصحاب المصالح والمسوح المصغرة، ومجموعات النقاش البؤرية/ المركزة وغيرها.	ما هو المطلوب سواء من حيث جمع البيانات والتحليلات؟
مقيم مستقل (يمكن أن يكون هو نفسه من يقوم بإجراء تقييم الأداء أو الأثر).	فريق تقييم مستقل، يتضمن على سبيل المثال مقيم رئيسي ومنسق ميداني وشركة دراسات.	مقيم داخلي أو مستقل.	عادة مقيم مستقل (ولكن يمكن أن يكون داخلياً أيضاً).	من الذي ينفذ التقييم؟
تقييم وتحليل اقتصادي لتكاليف البرنامج وفوائده.	تحليل إحصائي واقتصاد قياسي، مع إمكانية استخدام أساليب نوعية.	تحليل العمليات، باستخدام أساليب كمية ونوعية.	تحليل البرامج، مع إمكانية استخدام أساليب كمية بسيطة.	ما هي المهارات اللازمة؟
١٠,٠٠٠ دولار - ٣٠,٠٠٠ دولار	تتراوح التكلفة بين ١٥,٠٠٠ دولار إلى ١ مليون دولار أو أكثر، بناءً على حجم ومدى تعقد البرنامج.	١٠,٠٠٠ دولار - ٦٠,٠٠٠ دولار	١٠,٠٠٠ دولار - ٣٠,٠٠٠ دولار	ما هي تكاليف التقييم؟
- فاعلية التكاليف: كل برنامج - المنافع والتكاليف: نفس الشيء كما في تقييم الأثر	- البرامج التي تكون: - مبتكرة وغير مختبرة - ذات أهمية استراتيجية ومؤثرة - قابلة لإعادة التطبيق	كل برنامج	كل برنامج	ما هي البرامج الأكثر ملائمة لهذا التقييم؟

المصدر: مقتبس من Rubio (٢٠١١).

النقاط الرئيسية

١. تمثل أهداف التعلم الخاصة بنا نقطة الانطلاق لأي تقييم. ويتطلب ذلك صياغة أسئلة التقييم لكل مراحل سلسلة النتائج وتحديد أولوية الأكثر الأسئلة ذات الصلة الأكبر من بينها. يمكن لأسئلة التقييم، بوجه عام، أن تكون وصفية أو معيارية أو أن تكون أسئلة أسباب وآثار.
٢. يعتمد اختيار استراتيجية التقييم على أسئلة التقييم. إذ قد لا تحتاج المعلومات الوصفية بالكامل إلى تقييم حيث تكون المتابعة كافية. أما الأسئلة المعيارية، فتتم الإجابة عنها في الغالب من خلال تقييمات سير العملية أو تقييمات الأداء. وإذا كانت الأولوية لأسئلة الأسباب والآثار، فهناك حاجة إذاً إلى القيام بتقييمات الأثر. أما تحليل المنافع والتكاليف وتحليل فعالية التكاليف، فيجيبان على ما إذا كانت التكاليف المتضمنة في تدخل ما مبررة أم لا.
٣. إن تقييمات الأثر وحدها – تلك التي بإمكانها أن تنشئ واقعاً مضاداً صادقاً – تسمح لنا بأن نثبت ما إذا كان البرنامج ناجحاً وبتوليد معارف يمكن تعميمها فيما بعد خارج إطار التدخل نفسه. وهذا ما يجعلها مختلفة عن سائر أنواع التقييم ويجعل منها أداة أساسية في تقييم التدخلات المتعلقة بسبل عيش الشباب.
٤. يعتمد اختيار نوع التقييم المناسب على السياق التشغيلي. فمن المهم جداً فهم ما إذا كانت تكاليف كل نوع تقييم من مال وعاملين ووقت ملائمة لتدخل ما.
٥. بما أن تقييمات الأثر عادةً ما تكون نوع التقييم الأكثر استنفاداً للموارد، يجب تطبيقها بشكل انتقائي للإجابة عن أسئلة استراتيجية أو لتقييم التدخلات التجريبية المبتكرة لاختبار تدخلات تجريبية مبتكرة لم يتم إثباتها بعد، ولكنها واعدة.

دراسة حالة لـ NUSAF: تقرير إجراء تقييم الأثر من عدمه

أسئلة التقييم

تمثل هدف التعلم الأساسي لـ NUSAF في تقدير الأثر السببي للمشاركة في برامج تدريب مهنية على سبل العيش الاقتصادي والإندماج الاجتماعي. فكانت الأسئلة التي تهم NUSAF هي ما إذا كان برنامج فرص الشباب يساعد على

- زيادة عدد الشركات والأعمال التجارية التي انشئت
- تخفيض مستويات البطالة
- زيادة عدد ساعات العمل مقابل أجر
- تحسين الاندماج في المجتمع والحد من الصراعات
- الحد من الفقر
- زيادة الرفاهية النفسية والاجتماعية

نظراً لطبيعة الآثار والنتائج التي تغلب على هذه الأسئلة، كان تقييم الأثر هو الاختيار الأمثل لأسلوب التقييم.

كانت دراسة حالة NUSAF مهمة أيضاً بآثار البرنامج على منظمات التدريب المحلية. وحيث إن ذلك لا يمكن تحديده بسهولة من خلال تقييم الأثر، فقد تقرر أن ذلك سوف يكون جزءاً من عملية متابعة برنامج فرص الشباب.

السياق التشغيلي لـ NUSAF

نظراً لكون برنامج NUSAF ممولاً من قبل البنك الدولي بدعم قوي لتقييم الأثر من حكومة أوغندا، كان السياق التشغيلي مواتياً لتقييم الأثر.

- التوقيت: تم التخطيط لاستراتيجية التقييم منذ بداية البرنامج. وهذا ما سمح بالمرونة اللازمة للتخطيط لتقييم صارم للأثر.
- الموارد: أمكن تخصيص الموارد اللازمة وتم تعيين فريق مؤهل خارجي لإجراء التقييم.
- السياق السياسي: جعل التقييم أولوية منذ البداية يشجع حوار أصحاب المصالح ودعمهم.

سمات NUSAF التي تبرر إجراء تقييم الأثر

كان برنامج فرص الشباب برنامج منح نقدية ضخم مصمم ومنفذ من قبل حكومة أوغندا. وقد أوحى حجم وتأثير البرنامج، بالإضافة إلى توقع إعادة تنفيذ البرنامج في المستقبل، بأن تقييم البرنامج كان طريقة ممتازة لزيادة المعرفة المحلية والعالمية لبرامج التدريب القائمة على المنح النقدية. بالرغم من أن هذه الأنواع من البرامج يتم تنفيذها بشكل متزايد، إلا أنها غير مختبرة بشكل عام. وبالإضافة إلى ذلك، فإن حقيقة أن البرنامج قد نفذته الحكومة يشير أن مثل هذا البرنامج قابل للتوسع ويمكن تكرار تطبيقه في بلدان أخرى.

المصدر: طبقاً لـ ياتلمان ومارتينيز وفبالا (٢٠٠٩).

قراءات رئيسية

Imas, L., and Rist, R. 2009. *The Road to Results: Designing and Conducting Effective Development Evaluations*, Washington, DC: The World Bank. (Chapters 6 and 7 are relevant to this note.)

http://books.google.com/books?id=NEsg-BtinIsC&printsec=frontcover&source=gs_ge_summary_r&cad=0#v=onepage&q&f=false

Rubio, G. 2011. "The Design and Implementation of a Menu of Evaluations." PREM Notes, The Nuts and Bolts of M&E Systems, No. 6. Washington, DC: The World Bank.

<http://siteresources.worldbank.org/INTPOVERTY/Resources/335642-1276521901256/premnoteME6.pdf>

بنود

بند ٥: إثبات أثر البرنامج

يمثل التشكيك الحثيث قوة خلاقية لأن معظم الضرر
يسببه أشخاص مسرفون في الثقة بأنفسهم يعتقدون أنهم
يعرفون الجواب في حين أنهم لا يعرفون.

— ويليام إسترلي

النوايا

الحسنة ليست كافية. وكبدل عنها، نحتاج أن نعرف أننا بالفعل نحسن من حياة
الأشخاص ولا نسبب ضرراً لهم أكثر من النفع دون أن نكون على دراية بذلك.
يتوفر الإثبات من خلال تقييمات الأثر، والتي، بخلاف أنواع التقييم الأخرى، تقدم دليلاً علمياً
على فاعلية البرنامج.

في هذا البند، نتحرى السؤال الأساسي لتقييم الأثر: «كيف يمكننا التأكد أن التغييرات التي تطرأ في
النتائج التي نراها ناجمة عن تدخلنا؟» فقد أوضحنا أن قياس الأثر يتطلب تقدير ما الذي كان
يمكن أن يحدث في حالة عدم وجود البرنامج. يمكن إجراء هذه التقديرات من خلال تحديد مجموعة
مقارنة من خلال أساليب تقييم تجريبية أو شبه-تجريبية. كما نعرض أيضاً السبب وراء كون الأسلوبان
هما الأكثر شيوعاً - مقارنة المشاركين قبل وبعد التدخل ومقارنة المشاركين مع غير مشاركون تم
اختيارهم بشكل مقصود - لا يمكن أن يقدموا تقديرات موثوقة فيها لنجاح البرنامج.

تحدي الإسناد

تساعدنا تقييّمات الأثر في الإجابة عن أسئلة محددة للغاية بخصوص برنامجنا. وعلى النحو المتناول في بند ٤، فأنها تحاول الإجابة عما إذا كان التدخّل (السبب) يحسن النتائج بين المستفيدين (الأثر). على سبيل المثال:

- هل يساعد برنامج التدريب المهني الخاص بنا على زيادة دخول المتدربين؟
 - هل يزيد منهج تنظيم المشروعات في الوسط المدرسي من معدلات إتمام المرحلة الثانوية واهتمام الطلاب بالتعليم العالي؟
 - هل يعزز برنامج الإرشاد لانشاء المشاريع من تأسيس الشركات والأعمال التجارية والاستدامة؟
- يمكن أن يكون إثبات السببية بين أنشطة التدخّل والنتائج التي نلاحظها معقداً لأن العوامل الأخرى قد تؤثر أيضاً على النتائج التي نهتم بها. على سبيل المثال، أن نلاحظ ببساطة زيادة تأسيس الشركات والأعمال التجارية بعد إتمام تنفيذ برنامجنا الخاص بتنظيم المشروعات ليس دليلاً على نجاح برنامجنا لأن عوامل أخرى مثل الظروف أو اللوائح الاقتصادية المحلية الخاصة ببدء الأعمال قد تكون تحسنت أثناء تنفيذ برنامجنا وساهمت في تأسيس الأعمال. وبالمثل، ملاحظة انخفاض تأسيس أعمال تجارية بعد تدخّلنا لا يعني بالضرورة أن تدخّلنا قد تسبب في تراجع في بدء الأعمال، بالعكس فقد يعكس سوء الظروف الخارجية الأخرى.
- يتمثل الغرض من تقييّمات الأثر في التغلب على تحدي الإسناد هذا، وذلك بقياس إلى أي مدى ساهم برنامج معين، وهذا البرنامج فقط، في التغيير الذي يطرأ على النتائج موضع الاهتمام.

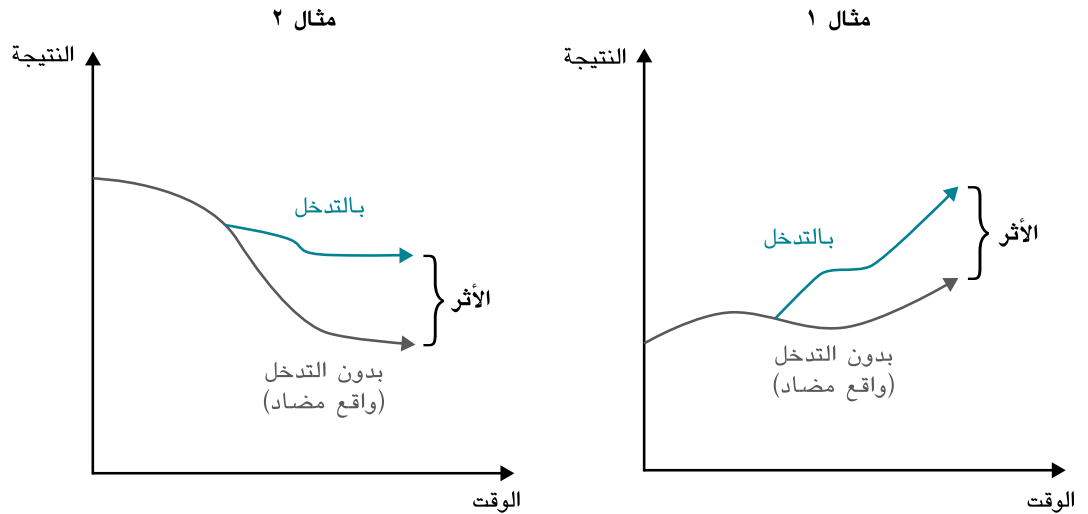
ما هو «الأثر» تحديداً؟

أولاً، نحتاج إلى توضيح ماذا نعني بكلمة أثر. يشير المصطلح غالباً إلى أهداف أو نتائج عالية المستوى لبرنامج تتعلق بتغييرات في مستويات المعيشة عموماً، مثل الحد من الفقر أو زيادة رفاهية الأفراد والأسر. وفي سياق تقييّمات الأثر، يتم فهم الأثر بشكل أكثر تحديداً على اعتبار أنه تغيير في النتائج والتي يمكن أن يعزى مباشرة إلى برنامجنا. ويكون التركيز هنا على «الإسناد المباشر»، والذي يعني أننا نريد أن نعرف أن التغييرات في النتائج التي نلاحظها هي في الواقع نتيجة لتدخلنا وليس لشيء آخر.

[تعريف]

النتائج مع البرنامج
– النتائج في غياب البرنامج
= الأثر

- ببساطة، على النحو المبين في شكل ١-٥، أثر التدخّل هو الفرق بين النتائج التي تم ملاحظتها مع التدخّل، و
- النتائج التي تمت ملاحظتها بالنسبة لنفس الفرد أو الأسرة أو المجتمع أو وحدة ملاحظة أخرى بدون التدخّل. والنتائج في غياب التدخّل هي ما نسميه الواقع المضاد، مشيراً إلى ما الذي كان يمكن أن يحدث للمستفيد إذا لم يتم تنفيذ البرنامج.



لأسباب واضحة، من المستحيل ملاحظة نفس الشخص (الأسر، المدرسة، وغيرها) بتدخل أو بدون. وعلى الرغم من أنه يمكننا ملاحظة التغيرات لهؤلاء الشباب الذين يشاركون في برنامجنا، من المستحيل أن نعلم وضعهم في غياب البرنامج. بمعنى آخر، لا يمكننا بشكل أكيد معرفة ما الذي كان يمكن أن يحدث لهم إذا لم يشاركوا في برنامجنا. ونتيجة لذلك، لن يمكننا أبداً الحصول على الواقع المضاد الحقيقي، لذلك يكون التقدير كافياً.

كيف يمكننا تقدير الواقع المضاد؟

لتقدير الواقع المضاد، نحدد مجموعات المقارنة والتي تعرف أحياناً بالمجموعات الضابطة. وتعرف مجموعة المشاركين في البرنامج بمجموعة المعالجة. تكون مجموعة المقارنة الجيدة لديها نفس خصائص مجموعة المعالجة باستثناء حقيقة أن أفراد مجموعة المقارنة لا يستفيدون من البرنامج. طبقاً لـ غيرتير وزملائه ٢٠١١ يجب أن تشترك مجموعتي المعالجة والمقارنة في نفس الخصائص بثلاث طرق على الأقل:

١. يجب أن يكونا متطابقين من ناحية خصائص قابلة للملاحظة وغير قابلة للملاحظة. تشير الخصائص القابلة للملاحظة إلى العمر والجنس ومستوى التعليم والوضع الاجتماعي والاقتصادي والخصائص الأسرية ووضع التوظيف وما شابه ذلك. تتضمن الخصائص غير القابلة للملاحظة الحافز والاهتمام والتفضيلات ومستوى الدعم العائلي وعوامل أخرى. على الرغم من أنه لا يجب أن يكون كل شخص في مجموعة المعالجة مطابقاً لكل شخص في مجموعة المقارنة إلا أن كلا المجموعتين يجب أن تكونا متماثلتين في المتوسط.
٢. يجب أن يكون من المتوقع أن تتفاعل مجموعتي المعالجة والمقارنة مع البرنامج بنفس الطريقة. على سبيل المثال، يجب أن يرجح زيادة نتائج مثل المهارات أو الدخل بالنسبة لأفراد مجموعة المعالجة بنفس قدر الزيادة بالنسبة لأفراد مجموعة المقارنة.
٣. يجب أن تكون مجموعتي المعالجة والمقارنة متعرضتين بقدر مساو للتدخلات الأخرى. على سبيل المثال، يجب أن تتاح لكلا المجموعتين نفس خدمات الدعم الأخرى المقدمة من الحكومة المحلية والمنظمات غير الحكومية وهكذا.

[تعريف]

مجموعة المقارنة هي مجموعة تشترك في نفس خصائص مجموعة المشاركين باستثناء حقيقة أن الأشخاص في مجموعة المقارنة لا يستفيدون من البرنامج. إن مصطلحي مجموعة المقارنة والمجموعة الضابطة يتم استخدامهما غالباً بالتبادل على الرغم من أن الأخير، على وجه التحديد، لا يكون قابلاً للتطبيق إلا في سياق التقييمات التجريبية (أنظر أدناه). من أجل غرض هذه الوثيقة سنستخدم المصطلح العام مجموعة المقارنة في الوثيقة بأكملها.

[تعريف]

تحيز الانتقاء يحدث عادة عندما يكون المشاركون وغير المشاركين في البرنامج مختلفين في خصائص لا يمكن ملاحظتها، والتي تؤثر على كل من قرار الفرد للمشاركة في البرنامج بالإضافة إلى النتائج موضع الاهتمام.

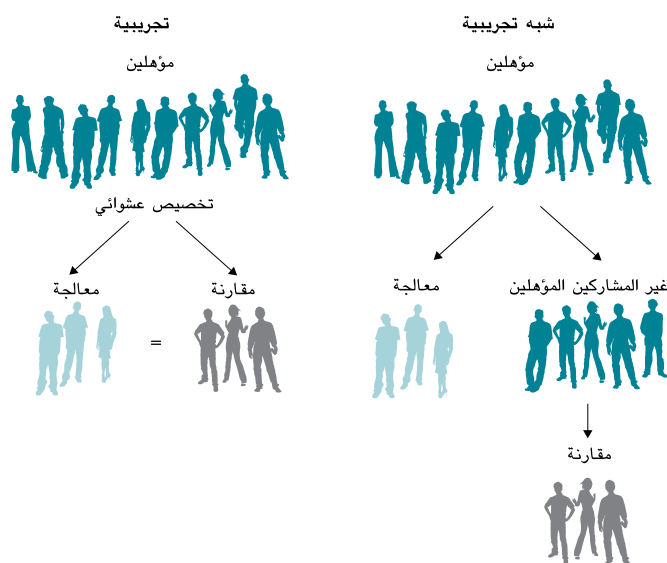
عندما تكون الشروط المذكورة أعلاه متساوية بين المجموعتين، لن يفسر وجود الاختلافات في النتائج سوى لوجود التدخل. في هذه الحالة، يمكن إثبات الأثر السببي للبرنامج. إذا ما كانت مجموعة المقارنة، من ناحية أخرى، مختلفة عن مجموعة المعالجة بطرق واضحة، سنواجه تحيز الانتقاء، مما سيجعل قياسات الأثر الخاصة بنا غير صادقة. يشير تحيز الانتقاء إلى حقيقة أن الاختلافات الكامنة بين مجموعتي المعالجة والمقارنة هي التي تفسر لنا سبب حصولنا على نتائج مختلفة، يظهر تحيز الانتقاء كثيراً عندما تتكون مجموعة المقارنة من أفراد إما غير مؤهلين للبرنامج (بناء على خصائص قابلة للملاحظة) أو ممن اختاروا عدم المشاركة (لأسباب غير قابلة للملاحظة).

في التدريب على المهارات وبرامج سبل العيش يرجح أن المتقدمين بطلبات المشاركة مختلفين عن غير المتقدمين بطلبات المشاركة، وأن هذه الاختلافات لا يمكن أن يراها الباحث بسهولة. على سبيل المثال، قد يكون المتقدمين بطلبات المشاركة محفزين بشكل أكبر من غير المتقدمين بطلبات المشاركة أو لديهم معلومات أفضل منهم. هذه الاختلافات قد تعني كذلك أن المتقدمين بطلبات المشاركة، في المتوسط، أنجح في سوق العمل من غير المتقدمين بطلبات المشاركة بغض النظر عن التدريب. في هذه الحالة، قد ترجع أسباب النجاح الأفضل بين متلقي التدريب إلى هذه الاختلافات الكامنة وليس إلى التدريب الذي تلقوه في البرنامج.

أساليب إيجاد مجموعات مقارنة جيدة

بوجه عام، توجد طريقتان لضمان أن مجموعتي المعالجة والمقارنة متشابهتين بقدر الإمكان: (١) بأساليب تجريبية و(٢) بأساليب شبه تجريبية (أنظر شكل ٢-٥).

شكل ٢-٥ الأساليب التجريبية في مقابل الأساليب شبه التجريبية



الأساليب التجريبية

تصميمات التقييم التجريبي ترتب عشوائياً من الذي سيكون في كل مجموعة بحيث إنه إذا كان لدينا مجموعة من المستفيدين المحتملين (مثلاً ٥٠٠ شاب، ٥٠٠ مدرسة أو غيره)، نختار عشوائياً بعضاً منهم (٢٥٠ على سبيل المثال) للحصول على البرنامج. هذه هي مجموعة المعالجة. أما الآخرين الذين لن يحصلوا على البرنامج؛ فيمثلون مجموعة المقارنة. إذا أجري الاختيار العشوائي بشكل صحيح سيرجح أن تكون كلا المجموعتين متشابهتين إلى حد بعيد (١) في الخصائص القابلة للملاحظة وغير القابلة للملاحظة، و(٢) في الطريقة التي يستجيبون بها للبرنامج، و(٣) في تعرضهم للتدخلات الأخرى. يشار إلى التقييمات التي تستخدم هذا الأسلوب، أو أشكال مختلفة منه، عادة بالتجارب المضبوطة العشوائية. أنظر مربع ١-٥ لمعرفة الاعتبارات الأخلاقية لتطبيق العشوائية.

يعزف بعض واضعي البرامج عن التخصيص العشوائي للمستفيدين المحتملين في مجموعتي المعالجة والمقارنة. ويتمثل مصدر القلق بصفة عامة في أن التقييم يؤدي إلى منع فوائد تبدو واضحة (مثل فرص التدريب) عن أفراد محتاجين وهو أمر غير أخلاقي. إلا أنه في الواقع من الخطأ افتراض أن المرء سيمنع ميزة ما إذا لم يكن البرنامج قد تم تقييمه بشكل صحيح. في البرامج التي لم يتم تقييمها قد يكون التخصيص العشوائي في الواقع أخلاقي أكثر من أساليب الاختيار الأخرى للأسباب التالية:

- **عدم التيقن من أثر البرنامج.** بالنسبة لمعظم البرامج، لا يكون واضحاً ما إذا كان البرنامج له أثر إيجابي على الفرد والمجتمع أو ما إذا كان هذا الأثر ذا حجم يبرر الموارد التي أنفقت. قد لا يكون للتدخل أي أثر أو حتى يكون له تأثيرات جانبية سلبية غير مقصودة. مثلاً البرامج المهيأة للفتيات مع استبعاد الصبية قد تزيد من العنف النوعي، وقد يتسبب برنامج تمويل المشاريع الصغيرة للشباب في ترك المشاركين بوضع أسوأ إذا لم يصبحوا قادرين على رد قروضهم. حتى برنامج التدريب، إذا تم تصميمه بشكل سيء، فقد يخفض فعلياً احتمالات الوظائف. عندما يتحقق أثر إيجابي (مثال، ١٠٠٠ دولار زيادة في الدخل لكل مشارك)، قد يأتي بتكلفة مرتفعة للغاية (مثال، ألف دولار للشخص)، مما يوحي بأنه كان من الأفضل بكثير إنفاق المال في موضع آخر، وعلى ذلك ففي حالة التدخلات التي لم يثبت بشكل كاف بعد هيكل الأثر والمنافع والتكاليف لها فسيكون تقييم البرنامج بناءً على مجموعتي المعالجة والمقارنة مبرراً بشكل جيد.

- **قيود الميزانية.** في الحقيقة، وبسبب الموارد المحدودة، يكون من النادر تحقيق إمكانية خدمة كل شخص محتاج حيث إن معظم البرامج لا توفر فوائد وخدمات إلا لعدد محدود من المستفيدين، وبذلك يستثنى آخرون، سواء كان ذلك بشكل صريح أم لا. على سبيل المثال، إذا كان لبرنامج تدريب للشباب عدداً محدوداً من الأماكن المتاحة سيتلقى بعض الشباب التدريب في حين أنه لن يتلقاه آخرون. وبالمثل، إذا أجري تدخل في منطقة محددة سيستثنى شباب مؤهل من المناطق الأخرى. يسمح التطبيق العشوائي للمسؤولين عن البرنامج باختيار مشاركين من المجتمع بطريقة عادلة وتعطي فرصة مشاركة مماثلة لكل شخص. إذا تم التطبيق العشوائي بشكل متفتح (على سبيل المثال مثل إجراء يانصيب في مناسبة عامة) فسيبرز كذلك الشفافية في عملية الاختيار وقد يقلل من مخاوف السكان من كون الاختيار قائم على تفضيلات شخصية أو سياسية.

من المهم كذلك ملاحظة أن التقييمات العشوائية لا تتطلب بالضرورة حرمان أي شخص من الخدمات. بند ٦ تقدم بيانات تفصيلية عن أساليب التقييم المختلفة.

الأساليب شبه التجريبية

إن تطبيق العشوائية لا يعد أمراً ذا جدوى أو مرغوباً على الدوام (أنظر مربع ٥-٢). في هذه الحالات قد تستخدم الأساليب شبه التجريبية لعزل تأثير تدخلنا. بالرغم من أنها عادة ما تكون أقل موثوقية من الأساليب التجريبية، إلا أن التصميمات شبه التجريبية تحاول محاكاة الواقع المضاد بتحديد غير المشاركين الذين يتشابهون مع مجموعة المعالجة بقدر الإمكان. وللقيام بذلك عادة ما تعتمد الأساليب شبه التجريبية على أدوات إحصائية وتحليل إحصائي. ويطلق على بعض الأساليب الشائعة تصميم عدم الاستمرارية، والاختلاف في الاختلاف، والمطابقة (أنظر بند ٦ لمناقشة تفصيلية).

مربع ٥-٢ أمثلة مختارة لحالات عدم إمكانية تطبيق العشوائية

- البرنامج قد بدأ بالفعل، المستفيدون تم اختيارهم بالفعل.
- الموارد المتاحة كافية لخدمة جميع الأفراد المؤهلين من مجتمع الدراسة. فقد يكون عندئذ منع الفوائد أو الخدمات فقط لغرض الدراسة أمراً غير أخلاقي.
- لا يمكننا اختيار مجموعة مقارنة أو استبعاد أي أحد من البرنامج. على سبيل المثال، فإن عمل حملة إعلامية لمحو الأمية المالية عبر التلفزيون أو الراديو يحتمل وصوله إلى كل أسرة ومن المستحيل متابعة من يستمع إليها ومن لا يستمع.
- يستهدف التدخل عدداً محدوداً من المجموعات أو المجتمعات ذات خصائص متفردة.
- هناك معارضة سياسية لتقديم التدخل لمجموعة دون الأخرى.

عندما تستوفى شروط مجموعة المقارنة الجيدة، نقول أن أثر التقييم قد تحقق له الصديق الداخلي (أنظر مربع ٥-٣).

مربع ٥-٣ الصديق الداخلي والخارجي

بشكل مثالي، تستوفي تقييمات الأثر متطلبين اثنين:

١. أن تنسم بالصديق الداخلي، ويعني أننا سنكون قادرين على إظهار السببية وللقيام بذلك يتم ضبط جميع الاختلافات الممكنة بين مجموعتي المعالجة والمقارنة بذلك سنكون قادرين على إسناد التغيرات بشكل واضح في النتائج للتدخل الذي قمنا به، ولضمان ذلك، نستخدم الأساليب التجريبية أو شبه التجريبية (تمت مناقشتها بالتفصيل في بند ٦).

٢. أن تنسم بالصديق الخارجي، ويعني أننا سنكون قادرين على تعميم النتائج. ذلك أنه يمكننا توقع نفس النتائج إذا قدمنا البرنامج إلى مجموعات مختلفة أو مجموعات أكبر، ولضمان ذلك، يلزمنا استراتيجية مناسبة لاختيار عينة الأفراد الذين نعمل معهم (ستتم مناقشة ذلك في بند ٧).

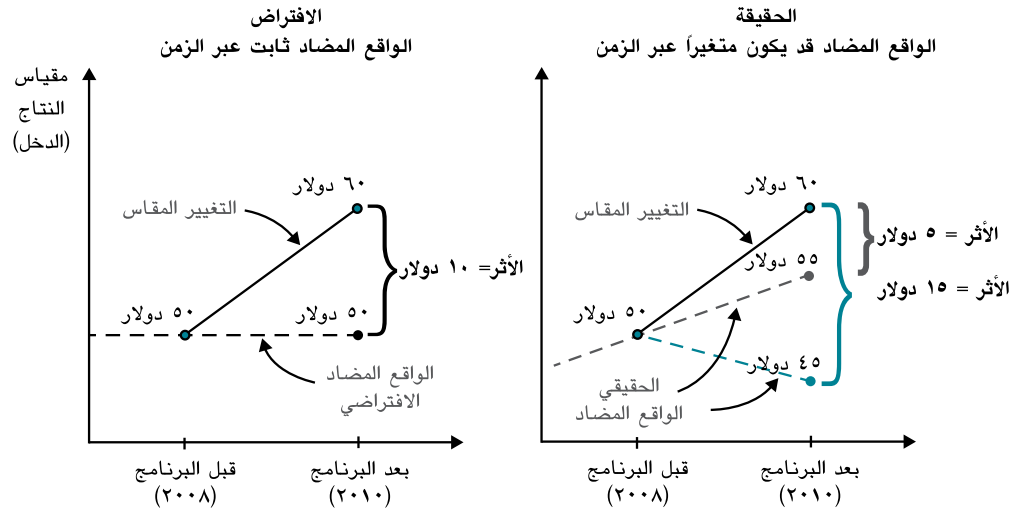
وقائع مضادة مزيفة

إن الأسلوبين الأكثر شيوعاً لقياس النجاح في برامجنا هما المقارنة القبلية والبعدية للتدخل بالنسبة للمشاركين، ومقارنة المشاركين بغير مشاركين يتم اختيارهم بموضوعية. وقد أخفقت هذه الأساليب في تحديد مجموعة مقارنة جيدة. وكنتيجة لذلك، لا يمكن اعتبارها أساليب مناسبة لتقييم الأثر فتقديراتها للأثر عادة ما تكون غير موثوقة، وفيما يلي السبب.

واقع مضاد مزيف ١: المقارنة القبلية والبعدية للمشاركين

في هذا الأسلوب، نستخدم نتائج ما قبل التدخل لتقدير الواقع المضاد. لذا نفترض أن البرنامج لو لم يوجد من الأساس فإن نتائج المشاركين بعد البرنامج ستكون مشابهة تماماً لمثيلتها قبل البرنامج. في مثال البرنامج التدريبي قد نلاحظ أن الدخل الشهري للمشاركين زاد من ٥٠ دولار قبل البرنامج إلى ٦٠ دولار بعد البرنامج، وبالتالي فقد نستنتج أن أثر البرنامج كان ١٠ دولار كل شهر لكل شخص (أنظر شكل ٥-٣، الرسم البياني على اليسار).

شكل ٣-٥ مخاطر في مقارنة النتائج القبلية والبعديّة



المشكلة

إن افتراض أنه في غياب البرنامج لم يكن ليتغير شيء ليس سوى افتراض غير مبرر في معظم الحالات. حيث يمكن أن يحدث العديد من الأشياء خلال فترة التنفيذ خاصة عندما تستمر البرامج لعدة سنوات. فعلى سبيل المثال، قد تحسن الأحوال الاقتصادية المحلية ويرتفع عدد الوظائف الشاغرة ويزيد معدل الدخول؛ وقد ترفع أحوال الطقوس الإيجابية كمية المحاصيل وتزيد الدخول في مجال الزراعة بالنسبة للعديد من الشباب أو تمكنت الحكومة من تطبيق برنامجها الخاص «العمل مقابل العائد» مما يرفع دخل العديد من الشباب. فإذا تحسنت، البيئة الخارجية تحسناً مستقلاً عن البرنامج لكان دخل الشباب ارتفع على أية حال (فلنقل، ٥٥ دولار في الشهر)، والأثر الحقيقي لتدخلنا سيكون على الأرجح أصغر من المقدّر بمقارنة بسيطة قبلية وبعديّة. في المثال الذي طرحناه سيكون المكسب ٥ دولارات بدلا من ١٠ دولارات (أنظر شكل ٣-٥، الرسم البياني على اليمين). وبالعكس، إذا ساءت الأحوال فعلياً (فلنقل أن الشاب قد يكسب ٤٥ دولار فقط في غياب البرنامج)، فإننا إذن قد نبخس قدر الأثر الحقيقي للبرنامج باستخدام مقارنة قبلية وبعديّة.

الخلاصة

هناك الكثير من العوامل التي يمكنها أن تؤثر على نتائج التدخلات المتعلقة بسبل عيش الشباب مع مرور الوقت. ونتيجة لذلك، لا يكون قياس النتائج قبل تنفيذ البرنامج في الغالب تقديراً جيداً للواقع المضاد. ولهذا السبب، لا تعتبر المقارنة القبلية والبعديّة أسلوباً جيداً لإثبات الأثر.

الواقع المضاد المزيف ٢: مقارنة المشاركين وغير المشاركين

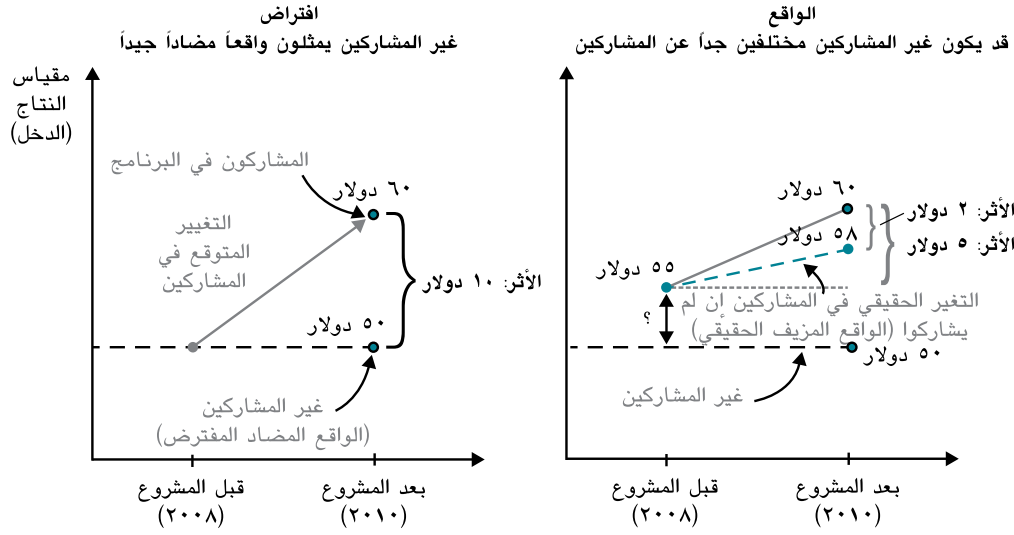
في هذا الأسلوب، نقوم بملاحظة نتائج غير المشاركين الذين يتم اختيارهم بشكل موضوعي في نهاية التدخل لتقدير الواقع المضاد. عند مقارنة المشاركين بغير المشاركين، فإننا نفترض أن هذه المجموعات متشابهة كثيراً في طبيعتها. فعلى سبيل المثال، نثق بأن كلتا المجموعتين تتشاركان في نفس الخصائص التي يمكن ملاحظتها والتي لا يمكن ملاحظتها، وأنهما ستستجيبان للبرنامج بنفس الطريقة وأنهما معرضتان بشكل متساو لتدخلات أخرى.

باستخدام مثالنا الذي يتمثل في برنامج تدريبي، فسوف نقوم بقياس مستوى دخل كل من المشاركين وغير المشاركين في نهاية التدريب. على فرض أن المشاركين يكسبون ٦٠ دولاراً شهرياً، بينما يكسب غير المشاركين ٥٠ دولاراً شهرياً، فقد نستنتج عندها أن أثر برنامجنا تمثل في ١٠ دولار شهرياً لكل شخص (أنظر شكل ٤-٥، الرسم البياني إلى اليسار).

[معلومة إرشادية]

تشمل الحالات المحددة التي يمكن فيها أن توفر المقارنات القبلية والبعديّة واقعاً مضاداً ثابتاً وحقيقياً، التدخلات المستهدفة قصيرة الأمد التي تهدف إلى تحسين اتجاهات ومعارف ومهارات محددة (أنظر، علي سبيل المثال، برنامج اعرف عن العمل التابع لمنظمة العمل الدولية في مربع ٥-٤ أدناه). في هذه الحالة، يمكن أن تعزى النتائج بشكل واقعي إلى أحد التدخلات المختارة في بعض الأحيان. ومع ذلك، تتأثر الآثار المحتملة الأخرى لبرنامج ما، مثل تغيير السلوك والتوظيف والدخل بالعديد من العوامل وبالتالي لا يمكن تقديرها بدقة بمجرد إجراء مقارنة قبلية وبعديّة.

شكل ٤-٥ مخاطر في مقارنة المشاركين بغير المشاركين



المشكلة

توجد مشكلتان أساسيتان تتعلقان بهذا النهج. المشكلة الأولى، إن افتراض أن المجموعتين لديهم مستويات متساوية من النتائج في بداية البرنامج قد لا يكون صحيحاً. قد يكون المشاركون، قبل البرنامج، أفضل أو أسوأ من غير المشاركين الذين يتم اختيارهم بشكل موضوعي. إذا قمنا بقياس النتائج فقط في نهاية البرنامج، فقد لا نتمكن من معرفة الظروف الأساسية. ربما يكون دخل المشاركين في بداية البرنامج أكبر من دخل غير المشاركين (على سبيل المثال، ٥٥ دولار) وبالتالي يكون التغيير الحقيقي مقارنة في ملاحظتنا في نهاية التدريب (٦٠ دولار) هو ٥ دولار بدلاً من ١٠ دولار (أنظر شكل ٤-٥، الرسم البياني على الجهة اليمنى).

ثانياً، عادة ما يكون افتراض أن المشاركين وغير المشاركين متشابهين جداً، غير صحيح. دعونا فقط نتأمل في المعايير التي نستخدمها في اختيار الشباب في البرنامج. ربما تكون المعايير مبنية على أساس من يأتي أولاً يحصل على الخدمة أولاً. وفي هذه الحالة، يكون الأشخاص الذين يتاح لهم الحصول على المعلومات حول وجود البرنامج هم الأشخاص الذين يقطنون بالقرب من مناطق تنفيذ البرنامج والذين يحصلون على التشجيع من والديهم أو ببساطة الأشخاص الذين لديهم حافز للمشاركة، ومن المحتمل في النهاية أن يشارك هؤلاء الأشخاص في البرنامج. ومن ناحية أخرى، فإن معايير الاختيار الواضحة مثل الدرجات التي يحصل عليها المرشحون في الاختبار أو المقابلات أو جودة خطة العمل، تشير إلى أننا صراحة نريد مشاركين يختلفون عن غير المشاركين. في كلتا الحالتين، وسواء كان هذا مرغوباً أو غير مرغوب، فمن المحتمل أن تكون هناك اختلافات في المتوسط بين المشاركين وغير المشاركين، وبالتالي، فإن المقارنة بين المجموعتين تكون مضللة. في الواقع، بالنظر إلى دوافع الشباب التي من المحتمل أن تكون أعلى، ووصولهم إلى المعلومات بشكل أفضل، وقربهم من موقع الخدمات، وغيرها - وهي صفات قد لا تكون مرئية دوماً بالنسبة لنا - ربما كان باستطاعة الشباب الذين شاركوا في برنامجنا أن يحسنوا أوضاعهم حتى دون التدخل. وبالعودة إلى المثال الذي ضربناه، إذا كان من الممكن للمشاركين كسب مبلغ ٥٨ دولار بعد فترة زمنية معينة حتى دون المشاركة في برنامجنا، فإن إجمالي العوائد التي يحصلون عليها بعد التدريب (٦٠ دولار) سيعكس أثر البرنامج بقيمة ٢ دولار فقط، وليس ١٠ دولار (أنظر شكل ٤-٥، الرسم البياني على الجهة اليمنى).

الخلاصة

عادة ما تكون هناك أسباب كامنة وراء مشاركة بعض الأشخاص في برنامج ما، في حين لا يكون لدى البعض الآخر أسباب لهذه المشاركة. هذه الأسباب تجعل من المشاركين وغير المشاركين أشخاصاً مختلفين في الواقع، سواء لاحظنا هذا الاختلاف (من خلال درجات الاختبار) أم لا (الدعم العائلي

والخافز). ونتيجة لذلك، في الغالب لا يمثل غير المشاركين الذين يتم اختيارهم بشكل موضوعي واقعاً مضاداً جيداً لفهم كيفية أداء المشاركين في ظل غياب البرنامج. وبالتالي، فإن لا يعتبر عقد مقارنة بسيطة بين المشاركين وغير المشاركين دون استخدام الأساليب التجريبية وشبه التجريبية أسلوباً جيداً لإثبات الأثر.

وعلى الرغم من أن الحقائق المضادة المزيفة المبينة أعلاه قد لا تكون مفيدة في تقدير الأثر - أي في الإجابة عن الأسئلة المتعلقة بالأسباب والآثار - إلا أنها قد لا تزال ذات قيمة بالنسبة لبرنامجنا. ففي الواقع، يمكن أن يكون جمع المعلومات الوصفية حول المشاركين وحتى حول غير المشاركين مهماً لإدارة البرنامج إذ أنه يمكن أن يساعدنا ذلك في فهم ديناميكيات برنامجنا بشكل أفضل. إنه من المشروع تماماً أن يتم استخدام هذه الأنواع من المقارنات كجزء من المتابعة أو تقييم الأداء الذي نجريه، طالما كنا مدركين ما الذي يمكن أن تخبرنا به نتائج هذه المقارنات وما الذي لا يمكنها أن تخبرنا به (للاطلاع على أمثلة، أنظر مربع ٥-٤).

مربع ٥-٤ أمثلة مختارة من التقييمات غير التجريبية

أسلوب المعالجة: المقارنة قبل وبعد البرنامج

برنامج اعرف عن العمل التابع لمنظمة العمل الدولية، سوريا

Assessing the Effect of Know About Business (KAB) on the Knowledge and Attitudes of Secondary School Students (2007)

http://www.syriatrust.org/site/images/files/KAB_Schools_Report_0708.pdf

أسلوب المعالجة: مقارنة المشاركين بغير المشاركين

منظمة إنجاز الشباب، الولايات المتحدة الأمريكية

The impact on students of participation in JA Worldwide: Selected cumulative and longitudinal findings (2004)

http://www.ja.org/files/long_summary.pdf

النقاط الرئيسية

١. الأثر الذي يحدثه برنامج ما هو التغيير في النتائج الذي يمكن أن يعزى مباشرة إلى التدخل. يتطلب فهم الأثر أن نقوم بعزل آثار البرنامج عن العوامل الأخرى التي تؤثر على نتائج المستفيدين.
٢. قياس أثر البرنامج يتطلب واقعاً مضاداً، مع معرفة ما الذي كان من الممكن أن يحدث للمشاركين في برنامجنا في حالة غياب التدخل.
٣. ومن أجل تقدير ما الذي كان من الممكن أن يحدث للمشاركين في حالة غياب البرنامج، نقوم بإنشاء مجموعات مقارنة تشارك بأكبر عدد ممكن من الصفات مع المستفيدين. إذا كان من الممكن تحديد مجموعة مقارنة جيدة، فإن مقارنة النتائج بين مجموعة المقارنة والمستفيدين (مجموعة المعالجة) ينتج عنها الأثر المترتب على البرنامج.
٤. ومن أجل العثور على مجموعات مقارنة جيدة، يمكن تصنيف أساليب تقييم الأثر في نوعين. تقسم الأساليب التجريبية مجتمع الدراسة المؤهل عشوائياً إلى أفراد يتلقون البرنامج وأشخاص لا يتلقونه. تحاول الأساليب شبه التجريبية العثور على مجموعة مقارنة جيدة من بين غير المشاركين، وذلك بمحاكاة مجموعة المعالجة إلى أقرب قدر ممكن.
٥. المقارنات القبلية والبعديّة البسيطة وكذلك مقارنة المشاركين بغير مشاركين يتم اختيارهم بشكل موضوعي مدروس لا توفر تقديرات موثوقة للأثر، فالأولى تفشل في ضبط التغييرات في العوامل الخارجية مع مرور الزمن والثانية تفشل في ضبط الصفات (غالباً لا يمكن ملاحظة ذلك) التي تؤثر على وضع البرنامج. ومع ذلك، يمكن أن تكون كلتاها مفيدتين في توفير المعلومات الوصفية كجزء من نظام المتابعة الخاص بنا.

دراسة حالة NUSAF: تحديد واقع مضاد

كان تحديد الواقع المضاد يشكل اهتماماً خاصاً لتقييم برنامج فرص الشباب. هل المشاركون في NUSAF يختلفون عن عامة السكان؟ إذا كان الأمر كذلك، فكيف يمكن اشتقاق واقع مضاد من خلالهم؟

توقعت الحكومة وفريق البحث أنه سيكون هناك اختلافات مهمة بين المشاركين في NUSAF وعامة السكان. وتمثلت إحدى الدلائل في أن الأفراد كان يفترض بهم أن يشكلوا مجموعات ويقدموا دراسة. وهذا يعني أنه كان يتعين على المتقدمين أن يكونوا متعلمين إلى حد ما، مما يعني أنهم أفضل حالاً. بالإضافة إلى ذلك، كان ينبغي على الذين تقدموا بمقترحات للبرنامج أن يكون لديهم رغبة في المشاركة في العمل، ولذلك فمن المحتمل أنهم كانوا متحفزين للغاية. وهذه الصفة لا يمكن قياسها بسهولة.

للتحقق من الاختلافات المحتملة، قام NUSAF بالإطلاع على المشاركين في البرنامج الذين تم اختيارهم في الأساس وقام الصندوق بمقارنة المشاركين مع شباب آخرين شملتهم عملية مسح تم إجراؤها في نفس الوقت تقريباً. وتم التوصل إلى أن أعضاء برنامج فرص الشباب كانوا يملكون أصول أكثر وكانوا على درجة من التعليم أعلى من عامة السكان. بالإضافة إلى ذلك، كان تمثيل المرأة في البرنامج منخفضاً للغاية (٣٣٪) مقارنة بعامة السكان (٥١٪). وكانت فرصة الأسر المشمولة في الدراسة لامتلاك جهاز راديو أو دراجة خمس مرات أكبر من فرصة عامة السكان، أما فرصتهم في الحصول على جهاز هاتف محمول أو ماشية أكبر بثلاث مرات من فرصة عامة السكان. كذلك كانت هناك فوارق في التعليم بين المشاركين في البرنامج وعامة السكان.

بالإضافة إلى ذلك، بمقارنة معدلات الفقر بين عامة السكان بمعدلات الفقر بين المشاركين في البرنامج، تم التوصل إلى فروق كبيرة: كان من بين المشاركين ٥٠٪ على الأقل ممن هم فوق المستويات المحددة للفقر. لذلك، سواء بالنظر إلى الفرق النسبي أو المطلق بين عامة السكان والمشاركين في برنامج فرص الشباب، كان من الواضح أن المتقدمين إلى البرنامج كانوا، في المتوسط، جزءاً من طبقات اجتماعية واقتصادية أعلى من عينة تمثيلية من الشباب في المنطقة.

لقد أكدت الاختلافات عبر المجموعات على أسباب ضرورة إجراء تقييم أفر دقيق. إن استخدام عامة السكان كواقع مضاد سوف يؤدي إلى مبالغة كبيرة في تأثير البرنامج، إذ أنه كانت هناك اختلافات كبيرة أصلاً في عينة مجتمع الدراسة حتى دون وجود البرنامج.

من أجل تحديد واقع مضاد يتميز بالصدق، تمكن فريق التقييم من الاستفادة من حقيقة أنه كان هناك طلب كبير على البرنامج، ولكن كان هناك القليل من الأموال المتبقية. لقد أدت المشاكل المتعلقة بتحديد إحدى مجموعة المقارنة المناسبة المبينة أعلاه، مع قلة الأموال اللازمة لضمان مشاركة كل شخص مؤهل، إلى اتخاذ قرار باللجوء إلى الأساليب العشوائية لاختيار مجموعة المقارنة.

المصدر: باتمان ومارتينيز وفيالا (٢٠٠٩).

قراءات رئيسية

Gertler, P., Martinez, S., Premand, P., Rawlings, L., and Vermeersch, C. 2011. *Impact Evaluation in Practice*. Washington, DC: The World Bank. See Chapter 3.
<http://www.worldbank.org/ieinpractice>

بنود

بند ٦: تحديد أسلوب مناسب لتقييم الأثر

الهدف

من هذا البند تزويد الممارسين بنظرة عامة على مختلف الأدوات المتاحة لتقييم الأثر وكذلك توجيههم إلى أكثر الأدوات ملائمة للبرامج المحددة. نقدم هنا حقيبة أدوات مكونة من ستة أساليب هي الأكثر شيوعاً لإجراء تقييم الأثر، والتي تتميز بكونها منظمة بقدرتها على تشكيل واقع مضاد بأدنى نسبة ممكنة من الانحياز. فلكل وسيلة مزاياها وعيوبها، ولا يقتصر أساس اختيار أسلوب تقييم الأثر على جودة الأسلوب من الناحية النظرية، بل يمتد ليشمل السياق التشغيلي للبرنامج. ومن ثم يجب إشراك مدراء البرنامج في عملية تصميم التقييم للتأكد من استجابة التقييم للاحتياجات ولسياق التدخل.

الاختيار من بين أساليب تقييم الأثر

تختلف كل وسيلة من وسائل تقييم الأثر عن غيرها فيما يتعلق بالظروف التي تحقق فيها أفضل النتائج، فعادة لا يتناسب كل تقييم مع سياق كافة البرنامج. وبناء عليه، تساهم خصائص وظروف كل برنامج في اختيارنا لأسلوب تقييم الأثر المقرر استخدامه. نحتاج تحديداً، كما يوضح غيرنر وزملاؤه (٢٠١١)، ص. ١٤٣ - ١٤٩) أخذ بعض الأمور بعين الاعتبار كالتوقيت والتغطية والاستهداف والموارد:

التوقيت

هل بدأت المشكلة بالفعل؟ تتمثل القضية الرئيسية هنا في ما إذا كان من الممكن تضمين تقييم الأثر مع تصميم البرنامج. وكما سنوضح بالتفصيل لاحقاً، عندما يتم تخطيط تقييم الأثر منذ بداية البرنامج، ترتفع جودة التقييم بشكل هائل لتمكينا من استخدام نطاق أوسع بكثير من أساليب التقييم.

التغطية

هل يغطي البرنامج جميع الأشخاص المؤهلين؟ بشكل مثالي، نحن نرغب في خدمة أي شاب يحتاج إلى الخدمة، ويعد هذا الأمر أكثر سهولة لدى بعض البرامج أكثر من غيرها. إذا لم يستنفذ البرنامج المقدم كماً هائلاً من الموارد (مثل فتح حسابات توفير للقصر) أو إذا تم تقديمه من خلال القنوات الإعلامية (حملات محو الأمية المالية من خلال الراديو والتلفاز) فإننا عندها لن نرغب - أو ربما لن نتمكن - من منع أحد الأشخاص من الاستفادة من التدخل. غير أننا في معظم الأحوال لا نملك المصادر الكافية لتقديم برامج سبل عيش الشباب إلى كل الأشخاص المؤهلين للاستفادة، مما يحملنا على اتخاذ قرارنا بشأن من هم الشباب المؤهلون الذين سيشاركون ومن لن تتسنى لهم فرصة المشاركة. على الرغم من أن عدم قدرتنا على توصيل البرنامج لكافة الشباب يعد أمراً مؤزقاً ومحبطاً من منظور وضع البرامج، يتيح عدد طلبات الالتحاق الهائلة فرصة تحديد مجموعة مقارنة وإجراء تقييمات ذات جودة لأثر البرنامج الخاص بنا.

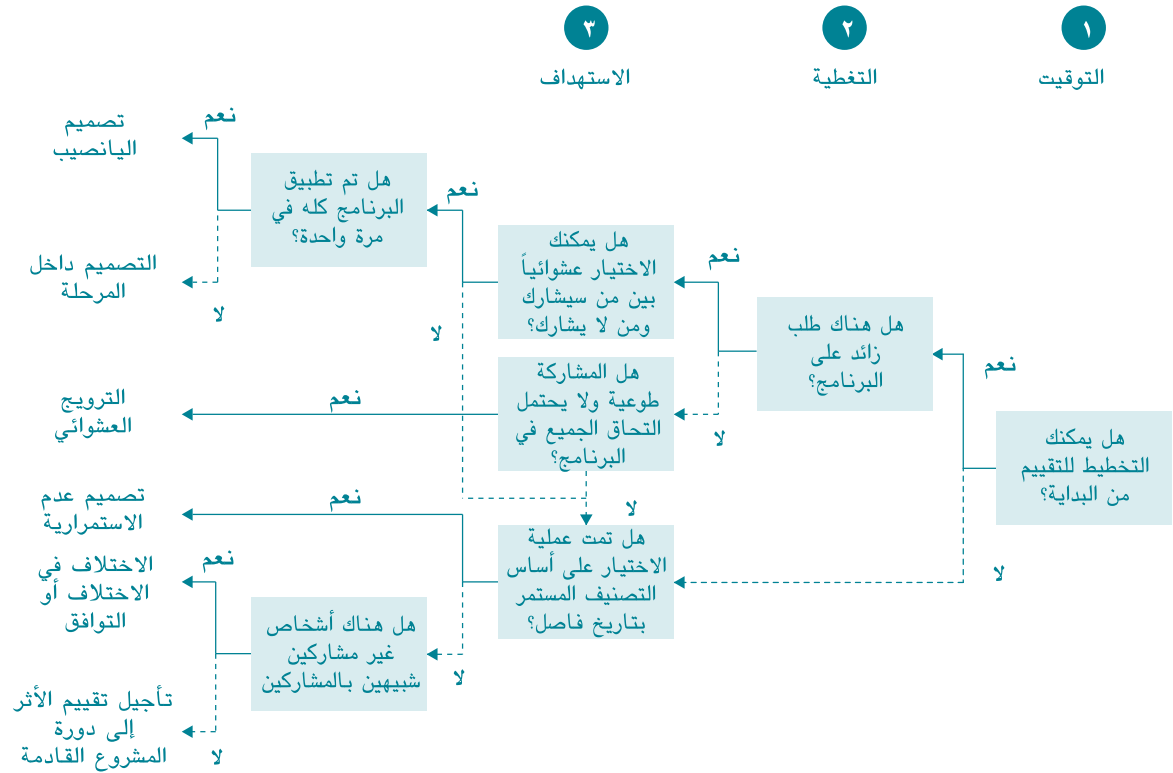
الاستهداف

كيف يختار برنامجنا المستفيدين منه؟ ما لم نتمكن من تقديم البرنامج إلى كافة الشباب المؤهلين، يتم اختيار المشاركين من الأفراد أو المجموعات من خلال الوسائل التالية:

١. **التخصيص العشوائي** هي العملية التي يتم من خلالها منح كل فرد أو مجموعة فرصاً متساوية للحصول على الفوائد. على سبيل المثال، سحب الأسماء من داخل القبة حتى يمكننا اتخاذ قرار من سيتسنى له فرصة المشاركة ومن سيردج في قائمة الانتظار.
٢. **تصنيف الأهلية** تحدد الأهلية وفقاً لعدد من المعايير الواضحة باستخدام حد واضح أو وضع حد أدنى نختار بناء عليه. ومن أمثلة تصنيف الأهلية تقديم المنح الدراسية اعتماداً على درجات الاختبار أو تقديم التدريبات بناء على مستويات الدخل.
٣. **قرار الاستهداف الاختياري** في بعض الأحيان لا تتوافر معايير واضحة تفسر اختيار أحد الأفراد أو المجموعات للمشاركة واستثناء آخرين، والذي يدل أن يضمن العدالة في الاختيار، فإنه يجعله منحازاً في اختياره للمشاركين. في حالات مثل ممارسات من يأتي أولاً يخدم أولاً والعوامل السياسية والأسباب الأكثر عملية من أساليب الاختيار المتحيزة بتأصل.

الموارد

هل يملك البرنامج الموارد اللازمة للقيام بتقييم الأثر المحدد؟ تختلف متطلبات وسائل تقييم الأثر فيما يتعلق بحجم العينة وجمع البيانات ودرجة تعقيد التحليل الإحصائي والتكلفة. حتى عندما نحدد الأسلوب الملائم للسياق التشغيلي لدينا، ربما يتم تنفيذه وربما لا حسب الموارد المتوفرة لدينا. يجب أن نأخذ الأسئلة الأربعة السابقة بعين الاعتبار أثناء التفكير في مختلف وسائل تقييم الأثر. وسوف نحدد الإجابة عن هذه الأسئلة آياً من الأساليب الستة تعد الأفضل في السياق الخاص بنا (أنظر شكل ٦-١). تتبع مناقشة لأساليب التقييم.



المصادر: تم صياغتها استناداً إلى دء (١٩٩١، ص. ٦٩): دفلو وغلينيستر وكريم (٢٠٠٦، ص. ٢٤-٢٧: غير تتر وآخرين (٢٠١١، ص. ١٤٨).

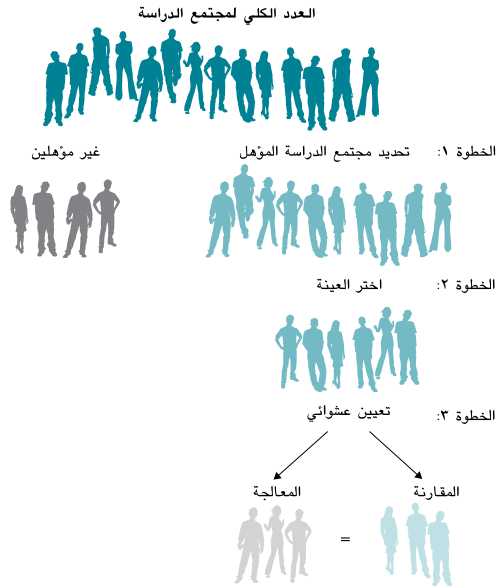
أسلوب ١: تصميم اليانصيب

اليانصيب هو طريقة بسيطة وشفافة لتخصيص الشباب للمجموعتين، التي سوف تحصل على خدماتنا (مجموعة المعالجة) والتي لن تحصل عليها (مجموعة المقارنة). يستخدم هذا الأسلوب من أجل تصميم تجارب عشوائية خاضعة للضبط. وتعد هذه الطريقة وسيلة تنظيم إحصائية بحيث إذا ما اختيرت عينة كبيرة نوعاً ما من نفس مجتمع الدراسة موضع الاهتمام، عشوائياً، لتخصيصها كإحدى المجموعتين، فإن المجموعتين في المتوسط سيكون لهما نفس الخصائص الملحوظة (العمر والنوع والطول والمستوى التعليمي وغيرها من هذا القبيل) بالإضافة إلى اشتراكهم في الخصائص غير الملحوظة (كالدافع والحالة العقلية). من خلال العشوائية، نجد أن اختلاف النتائج الملحوظة بين المجموعتين في نهاية برنامجنا قد ترجع إلى التدخل، وذلك يرجع لتساوي العوامل الأخرى التي قد تؤثر على النتائج في المتوسط. تعد تصميمات اليانصيب أكثر أنواع تقييم الأثر الحثيثة، ولذلك فإن النتائج عادة ما تحظى بثقة المانحين وأصحاب المصالح والحكومات.

كيفية العمل

يتم تصميم اليانصيب على ثلاث خطوات (أنظر شكل ٢-٦)

شكل ٢-٦ خطوات تصميم اليانصيب



الخطوة ١: حدد المجتمع المؤهل للدراسة

تحدد الخطوة الأولى من التجارب العشوائية الخاضعة للضبط في إيجاد مجموعة من الشباب المؤهلين للمشاركة في البرنامج. إذا أرادت إحدى عاملات الطب دراسة تأثير أحد العقاقير على مرض يصيب الأطفال، فإنها سوف تبحث عن مجموعة محددة من الأطفال، في حين أنها لن تسمح للبالغين أو المسنين بالإلتحاق في البرنامج. وبالمثل، فقد يستهدف برنامج سبل عيش الشباب المناطق الحضرية بأحد المراحل العمرية، وبذلك فلن يشتمل على بالغين أو شباب الريف. فمن المهم وضع معايير واضحة وشفافة (العمر والنوع ومستوى الدخل والوضع الوظيفي وغيرها) من أجل القدرة على تحديد من هم الأشخاص المؤهلين وغير المؤهلين للمشاركة.

[تعريف]

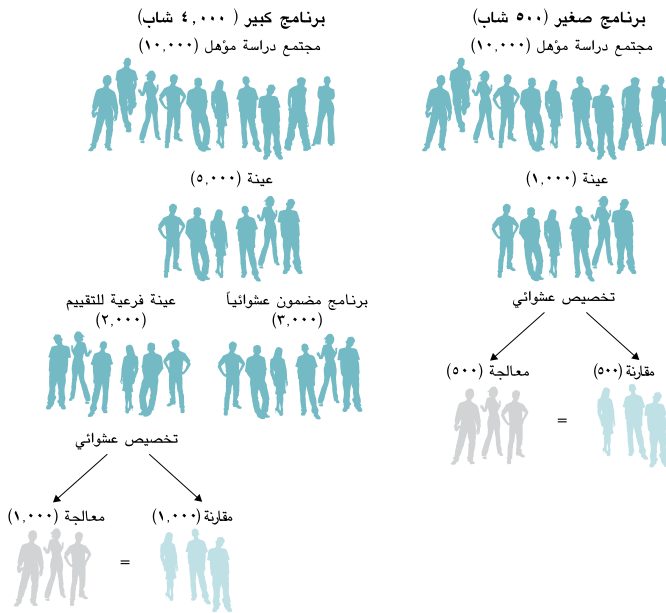
تجربة عشوائية خاضعة للضبط: هي دراسة يتم فيها توزيع الأشخاص عشوائياً (على أساس الصدفة وحدها) لتلقي المعالجة، مثل المشاركة في تدخل محدد.

العينة: جزء فرعي من مجتمع الدراسة. حيث تمثل عملية جمع المعلومات عن العدد الكلي لمجتمع الدراسة موضع الاهتمام مهمة مستحيلة أو غير عملية، يمكننا لذلك جمع معلومات عن قدر جزئي من مجتمع الدراسة يمكن إدارته. إذا ما أحسنا اختيار هذا الجزء، يمكننا عندها التوصل لاستنتاجات أو تقديرات استقرائية عن مجتمع الدراسة بالكامل.

الخطوة ٢: اختر عينة للتقييم

من أجل تقييم تدخّل ما، لسنّا في حاجة إلى اختبار كل من سيشارك في التدخّل. نحتاج فقط إلى اختيار مجموعة من الأشخاص الممثلين على أن يكون عددهم كافٍ ليفي بغرض التقييم الذي نحن بصدده؛ تسمى هذه المجموعة العينة (أنظر بند ٧) لمعرفة المزيد من التفاصيل حول كيفية تحديد العينة وحجمها). سيكون أفراد هذه العينة هم الشباب الذين سنعمل على جمع البيانات عنهم. يمكن اختيار عينة التقييم بطريقتين، وذلك حسب كبر أو صغر حجم البرنامج. فقد يجد البرنامج الصغير أن هناك ١٠,٠٠٠ من المستفيدين المؤهلين للانضمام إلى البرنامج، مثل شباب الشوارع في المناطق الحضرية ممن تتراوح أعمارهم بين ١٦-٢٤ عام. وقد تكون ميزانية البرنامج تكفي لمساعدة ٥٠٠ شاب منهم. والطريقة المثلى في هذه الحالة هو أن يتساوى حجم مجموعة المقارنة مع حجم مجموعة المعالجة، ومن ثم سيتم اختيار ١,٠٠٠ شاب من واقع ١٠,٠٠٠ من شباب الشارع للانضمام إلى البرنامج والتقييم (أنظر شكل ٦-٣، الصورة على الجانب الأيسر). قد تكون البرامج كبيرة الحجم أكبر من حجم العينة المطلوبة للتقييم. إذا كان البرنامج قادراً على خدمة ٤,٠٠٠ شاب، فليس من الضروري البحث عن ٤,٠٠٠ شاب آخر لإجراء المقارنة، لأننا لن نحتاج سوى ١,٠٠٠ فقط. ومن ثم يمكن للبرنامج تحديد عينة مكونة من ٥,٠٠٠ شاب من عدد مجتمع الدراسة الكلي الذي يبلغ ١٠,٠٠٠. ومن بين هذا العدد، يمكن ضمان انضمام ٣,٠٠٠ شاب إلى البرنامج. أما بالنسبة إلى العدد المتبقي وهو ٢,٠٠٠، فسيتم تقسيمه بشكل عشوائي بين البرنامج ومجموعة المقارنة (أنظر شكل ٦-٣، الصورة على الجانب الأيمن).

شكل ٦-٣ اختيار عينات البرامج الصغيرة والكبيرة



[معلومة إرشادية]

تتمثل إحدى طرق الحصول على عينة عشوائية من الشباب في الحصول على قائمة بإجمالي عدد شباب الشارع من التعداد السكاني أو سجلات تسجيل الناجين أو أية قاعدة بيانات أخرى من هذا القبيل ثم إجراء عملية اختيار عشوائية من هذه القائمة. إذا لم يكن من الممكن الحصول على هذه القائمة، يتم الحصول على عينة عشوائية من خلال استهداف عشوائي للمناطق التي ينتشر فيها شباب الشارع مثل المراكز الحضرية. إذا كان من المعروف أن الشباب يقضون الوقت في ٥٠ مركزاً مختلفاً من مراكز المدينة أو البلد، يتم اختيار المراكز عشوائياً ثم اختيار جزء من الشباب المنتشرين في هذه المراكز للمشاركة في الدراسة، الأمر الذي سينتج عنه اختيار للشباب بأقل درجة من الانحياز. سوف يناقش بند ٧ اختيار العينة بمزيد من التفصيل.

حتى يتسنى لنا اختيار العينة الممثلة من بين إجمالي عدد الشباب المؤهلين والذي يبلغ ١٠,٠٠٠ شاب من شباب الشارع، يجب اختيار العينة (سواء ١,٠٠٠ في الحالة الأولى أو ٥,٠٠٠ في الحالة الثانية) بشكل عشوائي من العدد المؤهل. ومن خلال الاختيار العشوائي، سوف يكون للمشاركين في البرنامج خصائص مماثلة لإجمالي العدد المؤهل. ورغم إن دراستنا تضم عدد محدود من الشباب، يمكن تعميم الأثر المحتمل للبرنامج على العدد الكلي لمجتمع الدراسة، والذي يصل في هذه الحالة إلى ١٠,٠٠٠ شاب.

الخطوة ٣: إجراء التخصيص العشوائي

تتمثل الخطوة التالية في تخصيص العينة التي تم اختيارها من الشباب لمجموعة المعالجة ومجموعة المقارنة والمتساويتين في الحجم تقريباً. في التجارب العشوائية الخاضعة للضبط، يكون لكل شاب فرصة في الانضمام إلى البرنامج. ويمكن أن يتم الاختيار العشوائي بالطرق التقليدية مثل قلب عملة معدنية أو رمي النرد أو سحب الأسماء من داخل القبعة. كما يمكن إجراء الاختيار العشوائي علنياً، في حالة الرغبة في ذلك وإذا كانت العينة صغيرة نسبياً (حيث أن سحب ٢,٠٠٠ اسم من داخل القبعة لا يبدو عملياً). وبدلاً من ذلك، وإذا كان عدد الأشخاص كبيراً يمكن إجراء الاختيار العشوائي عن طريق برنامج كمبيوتر مثل برنامج ميكروسوفت إكسل. يمكن إجراء الاختيار العشوائي على مستويات عدة (أنظر مربع ٦-١). من خلال التحديد العشوائي للعينة الخاصة بنا في مجموعة المعالجة أو المقارنة، نقوم باختيار المشاركين بشكل عادل كما نطور واقعاً مضاداً جيداً: فإذا كان حجم العينة كبيراً بشكل كافٍ، يكون للشباب في مجموعة المعالجة، في المتوسط، نفس الخصائص الملحوظة وغير الملحوظة للشباب في مجموعة المقارنة.

مربع ٦-١ مستويات العشوائية

يمكن إجراء التوزيع العشوائي على مستوى الفرد أو المجموعة أو المجتمع حسب احتياجات البرنامج.

مستوى الأفراد. تفيد العشوائية على مستوى الفرد أكثر في البرامج التي تقاس فيها النتائج على مستوى كل مشارك. قد تشوب هذه الطريقة بعض المشاكل مثل الدعايات التي تحدث عندما يتلقى الأفراد في مجموعة المقارنة بعضاً التأثير بطريقة غير رسمية، فالشباب الذين تلقوا تدريباً أو معلومات أخرى من برنامجنا مثلاً قد يشاركون معرفتهم أو مصادره مع أصدقائهم في مجموعة المقارنة.

مستوى المجموعات. إتباع العشوائية ليس مجدي أو مرغوباً عادة على مستوى الفرد. إن لم تتوفر قائمة بأسماء الأشخاص أو إذا كان يتوقع أن يتلقى الأفراد الذين تم اختيارهم في مجموعة المقارنة البرنامج على أية حال، قد تكون العشوائية على مستوى المجموعة في هذه الحالة أفضل. ويناسب ذلك بشكل خاص البرامج التي تعمل على مستوى المجموعات والتي تستهدف المدارس ومراكز التدريب المهني، ومراكز الشباب وأمثالها. في هذه الحالة، يتم اختيار مجموعات الأفراد عشوائياً للمعالجة أو المقارنة، فيتلقى كل الأفراد في مجموعة المعالجة التدخل نفسه. ويمكن للعشوائية على مستوى المجموعة أن تساعد على خفض الدعايات وقد تكون أسهل من العشوائية على مستوى الفرد. ومن ناحية أخرى، يمكن أيضاً إتباع العشوائية على مستوى المجموعة الفرعية، مثل الفصول في المدارس.

مستوى القرية/المجتمع. قد تختار البرامج أن تتبع العشوائية أيضاً على مستوى القرى والأحياء والمجمعات أو حتى المقاطعات، حين تطبق الأنشطة عند هذا المستوى، أو حتى حين يتوقع أن تتجاوز الدعايات مستوى المجموعة. فإذا كانت ١٠٠ قرية تهمننا في مقاطعة مثلاً ولا نملك الموارد الكافية لنعمل فيها كلها، فقد نختار عشوائياً العمل مع خمسين منها، فيما نبقى الخمسين الأخرى للمقارنة. ويكون جميع الشباب في القرى التي تحصل على المعالجة مؤهلين للمشاركة في البرنامج.

إن تطبيق التدخل على مستوى أعلى، وبالتالي إتباع العشوائية عند هذا المستوى، حتى وإن خفض من التأثيرات غير المرغوب فيها، يمكنه أيضاً أن يشكل مشكلة للأسباب التالية:

- كلما كان مستوى العشوائية أعلى، كلما أصبح عدد الملاحظات التي يمكن مقارنتها أقل. لذا يمكن التخفيف من حدة هذه المشكلة بإجراء مقابلات مع عدد من الأفراد في كل منطقة.
- يزداد حجم عينة التقييم مع ازدياد حجم مدى التدخل، ما قد يكون له تبعات فيما يتعلق بتكلفة التقييم.
- ومن المحتمل أن تختبر الوحدات الأعلى مستوى تأثيرات من الخارج بشكل أكثر مع مرور الوقت، الأمر الذي يضم تبعات لإمكانية المقارنة بين مجموعة المعالجة ومجموعة المقارنة، وبالتالي على الصدق الداخلي للتقييم.

لذا، يتعين على مدراء المشروع أن يجدوا مدى التدخل الأصغر الذي يمكن عنده تطبيق البرنامج وإتباع العشوائية فيه.

متى يمكنني استخدام تصميم اليانصيب؟

يستخدم تقييم اليانصيب العشوائي عندما يتم التخطيط للتقييم قبل التطبيق (المتوقع) وعندما يمكن للبرنامج أن يخدم قسماً صغيراً من الشباب المؤهلين. وما دامت القيود المرتبطة بالموارد تحول دون قيام البرنامج بخدمة كل السكان المؤهلين، فما من مشكلة أخلاقية في إبقاء مجموعة للمقارنة لأن قسماً من مجموع المجتمع المؤهل للدراسة سيظل خارج البرنامج بالتأكيد. في مثل هذه الحالة، يمكن إبقاء مجموعات المقارنة لقياس آثار البرنامج القصيرة والمتوسطة والطويلة الأجل (غير نلر وآخرون ٢٠١١). في أي تقييم متوقع، سيلزم جمع بيانات جديدة، وبالتالي ستتحتم تكاليف مرتبة. كحد أدنى، سيتوجب إجراء مسح ختامي (تم مناقشته بالتفصيل في بند ٧) للشباب في كلا من مجموعة المعالجة ومجموعة المقارنة. وفي كثير من الحالات، تدعو الحاجة لمسح أولي أيضاً. وبالرغم من التكاليف المرتبطة بجمع البيانات الجديدة، يمكن لليانصيب العشوائي البسيط أن يكون الخيار الأقل تكلفة لإجراء التقييم لأنه سيتطلب عدداً أقل من المسوح والمشاركين.

[تعريف]

التقييم المتوقع هو تقييم سيتم فيه متابعة المشاركين في المستقبل، لذا يجب التخطيط لهذا النوع من الدراسات في طور تصميم البرنامج.

أما التقييمات التي تنظر إلى المشاركين في البرامج التي قد تم تطبيقها أو تلك التي قد انتهت بالفعل، فيطلق عليها التقييمات بأثر رجعي.

المزايا

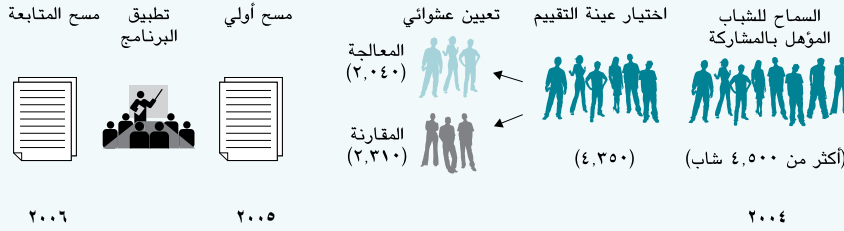
- إن تصميم اليانصيب هو الطريقة الأكثر صموداً لتطوير واقع مضاد، لأنه يقود إلى مجموعة مقارنة مطابقة جيداً (يستند إلى افتراضات أقل من الأساليب الأخرى). لذا، فهو يعتبر بالتالي التصميم الأكثر ثقة لقياس الأثر.
- وهو أبسط بكثير من الناحية التحليلية من جميع أساليب التقييم. ويكون أثر البرنامج في اختبار عشوائي هو بكل بساطة الفرق في النتائج بين مجموعتي المعالجة والمقارنة.
- وهو يتيح للجماعات أن تشارك بشكل مباشر في عملية الاختيار لتوزيع عادل وشفاف للمنافع.
- وبما أنه مخطط له منذ بداية البرنامج، فيمكن تصميمه لقياس متوسط أثر البرنامج وأيضاً ليقارن بين فاعلية مختلف عناصر البرنامج، ومختلف المدد الزمنية في البرنامج وغيرها من الأمور.
- سهل التطبيق والإبلاغ للعاملين بالبرنامج.

العيوب

- يتطلب أن يتم استبعاد مجموعة المقارنة من البرنامج طوال فترة تقييم الأثر.
 - يجب على المنظمات أن تضمن موافقة الشركاء وأصحاب المصالح المحليين على الطريقة.
 - يعتمد الصدق الداخلي لأسلوب تصميم اليانصيب على حقيقة أن العشوائية قابلة لتطبيق وأن يتم المحافظة عليها طوال الدراسة، الأمر الذي قد لا يكون من السهل تنفيذه. قد تنزعج صحة تطبيق هذا الشرط إذا لم يتم إتباع العشوائية بالشكل الصحيح أو إذا لم تكن مجموعة المعالجة أو المقارنة ممثلة لوضعها (أي إذا لم يكمل أفراد مجموعة المعالجة البرنامج أو إذا حصل أفراد مجموعة المقارنة على البرنامج)، أو إذا انسحب المشاركون من الدراسة قبل إتمامها أو إذا ظهرت التأثيرات من المشاركين على غير المشاركين.
- يقدم مربع ٦-٢ مثالاً لتصميم اليانصيب.

مربع ٦-٢ مثال لتصميم اليانصيب

استخدم أتانازيو وكوغلر ومغير (٢٠٠٩) تصميم اليانصيب لدراسة Jóvenes en Acción، برنامج لتوظيف الشباب في كولومبيا قدم ثلاثة أشهر من التدريب في الفصول وثلاثة أشهر من التدريب في مواقع العمل للشباب الذين تتراوح أعمارهم بين ١٨ و ٢٥ من الطبقة الاجتماعية الاقتصادية الأدنى. تم إعطاء تعليمات لموفري التدريب بأن يستقطبوا مرشحين أكثر مما كانت فصولهم تتسع له، في حالة عدم حضور الجميع إلى التدريب. ثم تم اختيار المشاركين عشوائياً من بين المرشحين المختارين، وتم وضع الشباب المتبقين على قائمة الانتظار واستخدامهم كمجموعة مقارنة.



كان أتانازيو وزملاؤه قلقون أنه بالرغم من إتباع العشوائية، قد تكون مجموعتي المعالجة والمقارنة مختلفتين بطرق لا يمكن للباحثين ضبطها. باستخدام البيانات الأساسية، نظروا إلى إمكانية المقارنة بين المجموعتين ووجدوا أنه، في المتوسط، كانت مجموعة المعالجة قد قصدت المدرسة قبل مجموعة المقارنة بثلاثة أشهر وكانت تضم نحو ٥ بالمائة من النساء الشابات أكثر من مجموعة المقارنة، غير أنهم لم يعتقدوا أن أيّاً من هذه الخصائص كان لها تأثير كبير على نتائج المعالجة.

كانت النتائج ككل واعدة. وفي المتوسط، كان أولئك الذين خضعوا للبرنامج مرجحين أكثر لأن يحصلوا على وظيفة رسمية مدفوعة وعلى رواتب أعلى وأن يحافظوا على وظائفهم لمدة أطول ممن في مجموعة المقارنة. وكانت الآثار أقوى بصفة عامة بالنسبة للنساء من الرجال.

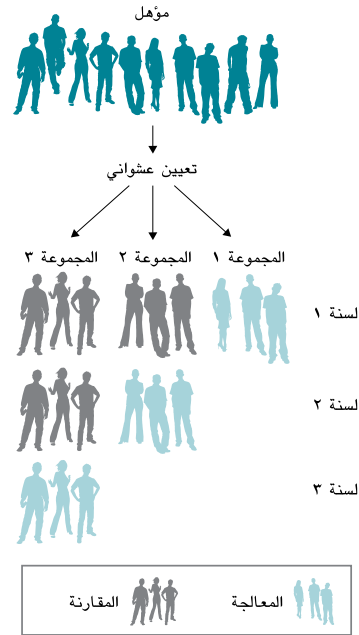
أسلوب ٢: تصميم عشوائي ينفذ على مراحل

يعتبر من المستحيل أحياناً إنشاء مجموعة مقارنة بحتة لم يتم فيها تقديم البرنامج للشباب أبداً. ولأن العديد من البرامج موجودة في مجتمع ما منذ سنين، فقد يكون عدم تقديم البرنامج أبداً لمجموعة من الأفراد المحتاجين أمراً صعباً من الناحيتين السياسية والمنهجية. إن التصميم المنفذ على مراحل هو شكل مختلف لتصميم اليانصيب. وهو ينطبق على البرامج التي يتم تقديمها على مر الزمن، ويستخدم تدفق المخرجات الطبيعي من أجل وضع مجموعات المعالجة والمقارنة.

كيفية العمل

يمثل الفرق الأساسي بين التصميم المنفذ على مراحل وتصميم اليانصيب في طريقة تخصيص الأفراد في مجموعات المعالجة ومجموعات المقارنة. وعندما يتم تقديم تدخل في عدة شرائح على مر الزمن، فإن التصميم المنفذ على مراحل يعطي كل شخص مؤهل أو مجموعة مؤهلة نفس الفرصة لتلقي البرنامج في إطار كل شريحة من الشرائح. ثم يتم بعدها اختيار مجموعة من الشباب عشوائياً من أجل تلقي المعالجة في الفترة الأولى، بينما يتم اختيار مجموعة أخرى لتلقي البرنامج في الفترة الثانية، ومجموعة ثالثة في الفترة الثالثة، وهكذا. وبالنسبة للوقت الذي توجد فيه مجموعات معينة على قائمة الانتظار، يمكن لهذه المجموعات أن تقوم بدور مجموعة مقارنة حتى تتلقى البرنامج (أنظر شكل ٦-٤).

شكل ٦-٤ مجموعات المعالجة والمقارنة في التصميم المنفذ على مراحل



ملاحظة: لا يتعين بالضرورة إيقاف المعالجة لكي يعمل التقييم، حيث سيستمر تنفيذ بعض التدخلات عندما يتم تطبيقها. ولكن يتم تقديم العديد من البرامج، مثل برامج التدريب، خلال فترة زمنية محددة.

على سبيل المثال، قد يوجد لدى منظمة غير حكومية ميزانية لتدريب ١,٥٠٠ شاب ولكن لا يكون لديها القدرة على إجراء كامل التدريب دفعة واحدة. وبدلاً من ذلك، فهي تختار تدريب ٥٠٠ شخص سنوياً لمدة ٣ سنوات. وإذا كان بإمكانها تحديد جميع المشاركين البالغ عددهم ١,٥٠٠ في البداية، فقد تكون العشوائية المنفذة على مراحل هي أفضل أسلوب تقييم لهم. حيث يتم تقسيم ١,٥٠٠ شاب بشكل عشوائي على ثلاث مجموعات. وفي السنة الأولى وبينما تتلقى المجموعة ١ التدريب، تنتظر المجموعتان ٢ و ٣ ويمكن أن تقوموا بدور مجموعة المقارنة. وفي السنة الثانية، تبقى المجموعة ٣ فقط لإجراء المقارنة. وبحلول السنة الثالثة، ستكون المجموعات الثلاثة قد حصلت على التدريب.

وبما أنه يتم اختيار الأفراد عشوائياً للمجموعات المختلفة، فإنه يمكن مقارنة هؤلاء الذين تلقوا المعالجة أولاً بهؤلاء الذين تلقوا المعالجة في وقت لاحق. ولكن وبما أن الجميع يحصل على البرنامج في نهاية المطاف، فإن التصميم المنفذ على مراحل لا يلائم عادة إيجاد الأثر طويل الأمد لبرنامج ما لأنه لا يوجد في النهاية مجموعة مقارنة. وحتى البرامج الكبيرة وطويلة الأمد ستواجه مصاعب في الطلب من المشاركين الانتظار لمدة ثلاث أو أربع سنوات، لذا فإن الفترة الزمنية للنتائج تقتصر عادة على سنة أو سنتين.

متى يمكنني استخدام التصميم المنفذ على مراحل؟

كما هو الحال في تصميم اليانصيب، فإن التقييم المنفذ على مراحل هو تقييم متوقع ويعتمد على الطلب المفرط والقدرة على تخصيص المشاركين عشوائياً لمجموعتي المعالجة والمقارنة. حيث إن التصميم المنفذ على مراحل أكثر ملائمة من تصميم اليانصيب في البرامج الكبيرة التي يتوقع فيها تقديم تدخلات على مدى سنين. ولأن التصميم المنفذ على مراحل يتطلب خطة موضوعة للتطبيق، فهو يتطلب أيضاً فريقاً مكرساً للبرنامج سيكون قادراً على متابعة التطبيق خلال دورة حياة البرنامج. لا تختلف التصميمات المنفذة على مراحل اختلافاً جوهرياً عن تصميم اليانصيب من ناحية البيانات أو متطلبات التكلفة. وسيكون هناك حاجة لإجراء مسح ختامي، وفي العديد من الحالات، هناك أيضاً حاجة لإجراء مسح أولي. وهناك فارق واحد مهم هو أن تكاليف تنفيذ البرنامج قد تزيد لأنه سيكون هناك حاجة للموارد من أجل ضمان تنفيذ التطبيق بالطريقة التي يتطلبها التقييم.

المزايا

- تعطي التصميمات المنفذة على مراحل واقعا مضاداً صامداً، وتكون عملية الاختيار فيها عادلة وشفافة، وتسمح بمقارنة آثار بدائل البرنامج.
- يلائم هذا الأسلوب التطبيق الطبيعي للعديد من البرامج.
- لأن الجميع في هذا الأسلوب يتلقى البرنامج في النهاية، فإنه يمكن للدراسات المنفذة على مراحل أن تكون ملائمة من الناحية السياسية.

العيوب

- كما في أسلوب اليانصيب، هناك تحديات لضمان العشوائية الناجحة والحفاظ على مجموعات المعالجة والمقارنة مع مرور الوقت.
- قد لا ينتظر المشاركون الانضمام للبرنامج. وإذا انتظروا، فإنه يوجد خطر أن يقوموا بتغيير سلوكياتهم في غضون ذلك الوقت وبهذا لن يكونوا قابلين للمقارنة في مجموعة المقارنة. فعلى سبيل المثال، قد يتوقفوا عن البحث عن عمل بانتظار الانضمام للبرنامج.
- على لا يستطيع الأسلوب المنفذ على مراحل تقدير الأثر طويل الأمد للبرنامج.
- يتطلب هذا الأسلوب إستراتيجية تطبيق واضحة، مما قد يؤدي لحدوث آثار تشغيلية.
- أنظر مربع ٦-٣ من أجل الإطلاع على مثال عن تصميم عشوائي منفذ على مراحل.

[معلومة إرشادية]

مع أسلوب التنفيذ على مراحل، يكون من المهم وجود الوقت الكافي بين كل مرحلة من المراحل لكي يعطي البرنامج آثاراً. وإذا كان مسؤول البرنامج يعتقد أن أثر البرنامج سيستغرق سنتين ليصبح نافذاً، فيجب أن يكون الوقت بين المرحلة الأولى والمرحلة الأخيرة سنتين على الأقل. وقد تكون البرامج الصغيرة أو قصيرة الأمد غير ملائمة لهذا الأسلوب.

يقدم برنامج البنك الدولي للتمكين الاقتصادي للفتيات المراهقات في ليبيريا تدريباً يستمر لمدة ستة أشهر ونشاطات متابعة تستمر أيضاً ستة أشهر مع منهجين مختلفين: (١) تدريب مهارات للعمل بأجر، مصحوباً بتقديم مساعدة في إيجاد وظائف، و(٢) مهارات تطوير العمل التجاري مصحوبة بالتشبيك للحصول على قروض صغيرة. كما يتم أيضاً توفير الإرشاد لجميع المستفيدات بدءاً من الشهر الثالث للتدريب.

ومن أجل تقييم آثاره، اختار البنك الدولي تصميم التقييم المنفذ على مراحل بما أن هذا التصميم سيسمح بإجراء تقييم عشوائي ذو جودة مع التمكن في نهاية المطاف من خدمة جميع الفتيات اللواتي وعدن بتلقي تدريب. وقد استفاد التقييم من التطبيق الطبيعي للبرنامج والقيود التشغيلية التي لم تسمح بتدريب الجميع في نفس الوقت.

وبعد المسح الأولي، تم التخصيص العشوائي لعدد ١,٢٧٣ مشاركة في مجموعة المعالجة (وتلقي التدريب خلال الجولة الأولى من البرنامج في عام ٢٠١٠) و٨٤٣ مشاركة لمجموعة المقارنة (وتلقي التدريب خلال الجولة الثانية من البرنامج في عام ٢٠١١). وتم إجراء مسح المتابعة في نهاية كل جولة وإتمامه باستطلاعات نوعية بعد التدريب لجمع المعلومات حول آراء المشاركات في التدريب الذي تلقينه والمحتوى وأصول التدريس والمدرسين.

ولأن البرنامج والتقييم استهدفا الفتيات اللواتي عبرن بالتحديد عن اهتمامهن بالتدريب، فإنه لا يمكن تعميم نتائج التقييم على أية شابة من مجتمع الدراسة. حيث يساعدنا التقييم على فهم أثر التدريب على اللواتي اخترن تلقي التدريب والمساعدة في الحصول على عمل بأجر أو تنظيم المشاريع.

المصادر: البنك الدولي (٢٠٠٨)؛ موزي (٢٠١١).

أسلوب ٣: التصميم العشوائي الترويجي

قد توجد حالات لا يمكن فيها إقصاء أي من المستفيدين المحتملين أو مثل هذا الإقصاء غير مرغوب إما بسبب أن المشاركة طوعية وإمكانية أي شخص الالتحاق إذا رغب بذلك، أو بسبب أن البرنامج يمتلك ميزانية كافية للخدمة الفورية لجميع المؤهلين من الشباب في مجتمع الدراسة. وفي مثل هذه الحالات، قد يكون أسلوب الترويج العشوائي (ويطلق عليه أيضاً التصميم التشجيعي) مناسباً.

كيفية العمل

يحدد الترويج العشوائي مجتمع الدراسة المؤهل ويختار عينة كما في تصميم اليانصيب أو التصميم المنفذ على مراحل. ولكنه يختلف من حيث العملية العشوائية. ففي حين لا يمكن التخصيص العشوائي للشباب لمجموعة ما تتلقى فوائد ومجموعة أخرى لا تتلقى فوائد، قد يكون من الممكن بدلاً من ذلك الترويج العشوائي على البرنامج. أي بدلاً من الاختيار العشوائي للذين يحصلون على الفوائد والخدمات، نقوم بالاختيار العشوائي للأشخاص الذين يتم تشجيعهم على الحصول على هذه الفوائد. يقوم التشجيع العشوائي على أساس أنه بالنسبة للعديد من البرامج، ستكون هناك ثلاث مجموعات من المستفيدين المحتملين:

- الشباب الذين لن يلتحقوا أبداً.
- الشباب الذين يلتحقون دائماً؛
- الشباب الذين يلتحقون فقط في حالة تم تشجيعهم على القيام بذلك.

بغض النظر عما يقدمه البرنامج، سواء كان حسابات توفير مجانية، أو تدريب مهني، أو برامج محو الأمية المالية القائمة على وسائل الإعلام، فإنه غالباً من غير المرجح أن يرغب كل شاب مؤهل في المشاركة. حيث قد لا يثق البعض بالتدخل ببساطة، وقد يواجه آخرون عوائق مثل الوقت أو وسائل المواصلات، وقد لا يعلم آخرون بالبرنامج ببساطة.

ببساطة يمكن أن يأخذ التشجيع العشوائي عدة صيغ. ففي حالة حسابات توفير الشباب، يمكننا التشجيع العشوائي من المبادرة في مدارس متتارة. وبالنسبة لبرنامج التدريب، فقد نستخدم أخصائي اجتماعي لكي يقوم بزيارات عشوائية لمنازل الشباب العاطلين عن العمل، ويقوم بوصف البرنامج ويعرض التحاق الشباب في الحال. وفي حالة حملة محو الأمية المالية، قد نرغب بإرسال رسائل نصية بشكل عشوائي لجزء من الجمهور المستهدف، ولكن ليس لآخرين. وفي جميع الحالات، سيظل هناك

أشخاص في المجموعة التي تم تشجيعها لن يحصلوا على البرنامج، مثلما سيكون هناك أشخاص في المجموعة التي لم يتم تشجيعها من سيحصلون على البرنامج بالفعل. ولكن الفكرة هي أنه إذا كان التشجيع فاعلاً، عندها يجب أن يكون معدل الالتحاق في المجموعة التي تم تشجيعها أعلى منه بين الأشخاص الذين لم يتلقوا التشجيع. وإذا تم التشجيع عشوائياً، عندها تشارك المجموعات التي تم تشجيعها والتي لم يتم تشجيعها في نفس الخصائص بالمتوسط، مما يسمح بتحديد الأثر السببي. للأسف، لا يمكننا فقط مقارنة نتائج من شاركوا في البرنامج بنتائج من لم يشاركوا. وكما تم نقاشه في بند ٥، فإن الأشخاص الذين يختارون المشاركة في برنامج ما يختلفون في أغلب الأوقات عن الذين لم يشاركوا، ويمكن أن تكون الكثير من هذه الاختلافات لا يمكن رؤيتها أو قياسها. وحتى إذا كان الترويج عشوائياً، فإن المشاركة في البرنامج لن تكون عشوائية، لذا فإن مقارنة المشاركين بغير المشاركين تكون أشبه بمقارنة التفاح بالبرتقال.

ولكن ما يمكننا فعله هو مقارنة نتائج جميع الشباب الذين تلقوا التشجيع بنتائج الشباب الذين لم يتلقوا التشجيع (أنظر شكل ٦-٥). دعونا نتأمل مثلاً عن برنامج للتدريب المهني حيث شارك في التدريب ٣٠ بالمائة من الشباب المؤهل في المجموعة التي لم يتم تشجيعها و٨٠ بالمائة من الشباب المؤهل في المجموعة التي تم تشجيعها (غير تملر وآخرون ٢٠١١). وبعد مرور سنة على البرنامج، لاحظنا من متوسط الدخل الشهري للمجموعة التي لم يتم تشجيعها كان ٦٠ دولار وللمجموعة التي تم تشجيعها ١٠٠ دولار.

قد يكون تقييم الترويج العشوائي مناسباً لـ:

- البرامج التي توزع قسائم للتدريب.
- البرامج التي تشجع الشباب على فتح حسابات توفير.
- التدخلات التي تقوم على حملات وسائل الإعلام الجماهيري.

شكل ٦-٥ تقدير الأثر طبقاً لتصميم التشجيع العشوائي

التغيير الملاحظ	مجموعة لم يتم تشجيعها	مجموعة تم تشجيعها	الالتحاق (٪) من مجتمع الدراسة (المؤهل)
٥٠٪	٨٠٪	٣٠٪	النوع ١: لا يلتحقون أبداً
			النوع ٢: يلتحقون دائماً
			النوع ٣: يلتحقون فقط في حال تشجيعهم
متوسط النتيجة (الدخل الشهري)	٦٠ دولار	١٠٠ دولار	٤٠ دولار
الأثر السببي			٨٠ دولار (٤٠ دولار/٠.٥)

أولئك الملتحقون بالفعل في كل سيناريو

المصدر: مقتبس من غير تملر وآخرين (٢٠١١، ص. ٧٥).

نظراً لأن تحديد الترويج يتم عشوائياً، فإن للمجموعات التي تم تشجيعها والتي لم يتم تشجيعها خصائص متساوية. وبالتالي، يمكننا أن نعزو الفرق الذي نلاحظه في متوسط النتائج بين المجموعتين (٤٠ دولار) إلى حقيقة أن مجموعة الأفراد الذين التحقوا فقط عندما تم تشجيعهم حصلوا على البرنامج. وعلى الرغم من أنه لا يمكننا مباشرة التفريق بينهم وبين الذين يلتحقون بشكل دائم، إلا أننا نعرف أن حصتهم من كامل مجتمع الدراسة هو الفرق في معدلات الالتحاق (٥٠ بالمائة، أو ٠,٥). وبالتالي فإن متوسط أثر البرنامج على المشاركين بسبب التشجيع هو ٨٠=٠,٥/٤٠ دولار.

متى يمكنني استخدام الترويج العشوائي؟

يلتزم الترويج العشوائي تماماً التقييمات المتوقعة للبرامج التي تكون ذات أهلية شاملة أو البرامج التي لا نستطيع فيها ضبط من يشارك ومن لا يشارك. وهو يعمل بشكل أفضل عندما يمكن لأحد أنواع التشجيع أن تؤثر بشكل كبير على أخذ البرنامج. ولا يعتبر الترويج العشوائي خياراً جيداً للخدمات التي تغطي شعبية كبيرة، مثل المنح النقدية، حيث يرغب الجميع في الحصول عليها عندما يسمعون عنها.

باستخدام هذا الأسلوب نحسب متوسط أثر برنامجنا بناءً على الأشخاص الذين انضموا للبرنامج نتيجة للجهود التشجيعية، ولأن هؤلاء المشاركون ليسوا إلا مجموعة جزئية من مجتمع الدراسة المؤهلين، يلزمنا عادة عينات كبيرة للغاية لهذا النوع من التقييم للتأكد من أن نتائجنا ذات دلالة إحصائية مما يزيد من عبء جمع البيانات. إذا تم الترويج على مستوى المجتمع قد يلزمنا مسح أشخاص من المجتمع أكثر بكثير مما كان سيلزمنا عند إجراء تصميم يانصيب بسيط. نتيجة لذلك، ستكون تكاليفنا على الأرجح أعلى من التكاليف المصاحبة لأنواع أخرى من التقييمات. الشروط الأخرى مبينة في مربع ٤-٦.

مربع ٤-٦: شروط ضرورية لينتج عن تصميم التشجيع

تقديرات أثر صحيحة

يجب أن يكون للمجموعة المشجعة والمجموعة غير المشجعة خصائص قابلة للمقارنة. يمكن تحقيق ذلك عن طريق التوعية العشوائية أو أنشطة الترويج للأفراد أو المجموعات أو المجتمعات في عينة التقييم.

يجب أن تزيد حملة الترويج من تسجيل من هم في المجموعة المشجعة بنسبة أعلى نسبياً من المجموعة غير المشجعة بشكل ملحوظ. إن «شكل ملحوظ» مفهوم نسبي قائم على احتياجات القوة الإحصائية. بوجه عام، يجب أن يزيد البرنامج المشاركة بنسبة ٤٠٪ أو أكثر حتى يصبح فاعلاً من ناحية التكلفة، ويمكن إثبات ذلك بالتحقق من أن نسب التسجيل أعلى في المجموعة التي تلقت التشجيع من المجموعة التي لم تتلقاه.

من المهم ألا يؤثر التشجيع نفسه مباشرة على النتائج موضع الاهتمام. إذا أدى الترويج إلى تغير السلوك فمن غير الممكن تحديد إذا ما كانت التغيرات التي لوحظت على الأشخاص بسبب البرنامج أو بسبب التشجيع، وأغلب الظن أن هذا سيحدث إذا تم التشجيع بالاقتران ببرامج التدريب. في معظم الحالات، تصبح الأهمية البالغة لمعرفة تأثير البرنامج وليس التشجيع.

المصدر: مقتبس من غيرتزر وآخرين (٢٠١١، ص. ٧٣).

المزايا

- لا تخرم حملات الترويج العشوائي أي شخص على الإطلاق من البرنامج بل أنها تسمح للأشخاص باتخاذ قراراتهم الخاصة حول ما إذا كانوا سيحصلون على البرنامج أم لا.
- ينتج هذا النوع من التقييم مجموعة مقارنة عالية الجودة بنفس خصائص، في المتوسط، مجموعة المعالجة تماماً مثل أي من أساليب التطبيق العشوائية الأخرى المذكورة أعلاه.

[تعريف]

في مجال الإحصاء يطلق على النتيجة ذات دلالة إحصائية إذا كان من غير المحتمل أن تكون هذه النتيجة قد حدثت بالصدفة. ولا تدلنا الدلالة الإحصائية على مقدار حجم التأثير (أهمية اقتصادية): أي أن أثر البرنامج قد يكون مهماً إحصائياً ولكنه صغير جداً.

العيوب

- هذا الأسلوب لا يمكن استخدامه إلا لبرامج معينة.
- غالباً ما يحتاج إلى أحجام عينات أكبر مما تحتاجه الأساليب الأخرى مما يزيد التكلفة.
- الأساليب الإحصائية المتقدمة مطلوبة لحساب أثر البرنامج.
- يجب أن يكون الباحثون حريصون عند تفسيرهم للنتائج لأن تقدير الأثر سيكون صحيحاً فقط بالنسبة للذين شاركوا في البرنامج لأنهم تم تشجيعهم؛ لا يمكن تعميم النتائج على المجموعات الأخرى من المستفيدين المحتملين.

مثال على تصميم عشوائي تشجيعي في مربع ٦-٥.

مربع ٦-٥ مثال لتصميم عشوائي ترويجي

في جنوب أفريقيا، تم استخدام تصميم عشوائي ترويجي لتقييم أثر التعليم الترفيهي الذي يهدف إلى تعزيز المعارف والمواقف والسلوكيات فيما يخص صنع قرار مالي محكم بتركيز خاص على إدارة الديون. يتكون البرنامج من تضمين سيناريوهات القدرة المالية في التمثيلية التلفزيونية التي تعالج مشاكل الحياة المنزلية الفضيحة، والذي استمر عرضها لسنوات عديدة.

يعد تقييم أثر التمثيلية التلفزيونية على السلوكيات والاتجاهات أمراً مثيراً للتحدي. أولاً، من الصعب فصل تأثير رسالة التمثيلية التلفزيونية عن الرسائل الأخرى الخاصة بموضوعات مماثلة قد يتلقاها الأفراد والأسر من مصادر أخرى. ثانياً، قد تختار أنواع معينة من الأفراد اختياريًا ذاتياً مشاهدة التمثيلية التلفزيونية تلك، وبذلك سيكون أي تغير يعقب ذلك في السلوك متداخل بسبب هذه المواقف الاختيارية. ثالثاً، حيث إن مشاهدة التلفاز أمر عالمي في الأساس، فمن الصعب تحديد مجموعة مقارنة جيدة من الأفراد الذين لا يتلقون رسائل القدرة المالية.

للتغلب على هذه المشكلات، تم تصميم منهجية الترويج العشوائي: بعد تحديد مجتمع الدراسة (تقريباً ١,٠٠٠ شخص)، تم تقديم حافظ مالي لنحو نصف مجتمع الدراسة (حوالي ١٠ دولارات) لمشاهدة الفضيحة؛ كانت هذه هي مجموعة المعالجة المختارة عشوائياً. تم التشجيع على مشاهدة البرنامج من خلال المكالمات قبل إجمالي ثلاثة عروض في فترة ثلاثة أشهر منبهين الأفراد بالحافز المالي لمشاهدة هذا العرض المحدد. أثناء هذه المكالمات علم أعضاء مجموعة المعالجة الظروف التي سيتلقون فيها حافظهم وتم توجيه عدد من الأسئلة لهم لتحديد المعرفة المسبقة عن الأمور المالية. بعد إذاعة العرض تم استدعاء الأفراد ومنحهم الحافز إذا أجابوا عن عدة أسئلة عن المحتوى غير المالي للعرض بشكل صحيح. أثناء نفس المكالمة، تم سؤالهم عدا من الأسئلة عن المعارف والاتجاهات المالية.

تم توفير نفس الحافز المالي للنصف الآخر من السكان - مجموعة المقارنة المختارة عشوائياً - لمشاهدة تمثيلية تلفزيونية مماثلة تمت إذاعتها تقريباً في نفس الوقت والأهم ليس بها محتوى محو أمية مالي. إن آلية منح الحافز كانت مطابقة لمجموعة المعالجة. سئلت مجموعة المقارنة نفس الأسئلة عن محو الأمية المالي مثل مجموعة المعالجة.

كانت النظرية، أنه إذا نجح محتوى التعليم المالي في الفضيحة؛ سيحزن الذين تم تشجيعهم على مشاهدة العرض درجات أفضل في الأسئلة المالية من المجموعة التي تم تشجيعها على مشاهدة التمثيلية التلفزيونية الأخرى. تم التقاط التأثيرات الفورية على المعارف والاتجاهات من خلال مسح قصير بعد نهاية العرض؛ تم التقاط تأثيرات طويلة المدى من خلال مسح متابعة متعددة.

المصدر: البنك الدولي (٢٠١١).

أسلوب ٤: تصميم عدم الاستمرارية

الحقيقة أنه في حالات عديدة لم تكن قادرين على تخطيط التقييم أثناء تصميم البرنامج، وحتى عندما تصبح قادرين على ذلك، فقد يكون من المستحيل استخدام أي شكل من أشكال تطبيق العشوائية للحصول على واقع مضاد صحيح. في هذه الحالات، قد نكون قادرين على استخدام قواعد استهدافية أخرى للبرنامج للحصول على مجموعة مقارنة جيدة. في الواقع تستخدم العديد من البرامج تصنيفاً مستمراً للمستفيدين المحتملين، مثل درجات الاختبار أو درجات الائتمان أو مؤشر الفقر أو العمر، ويكون لديها مستوى حد في قبول الدخول للبرنامج. على سبيل المثال، مقدمي الطلبات في مسابقة من مسابقات خطة الأعمال أو في بنك من بنوك تمويل المشاريع الصغيرة قد يتم إعطاؤهم درجات بناء

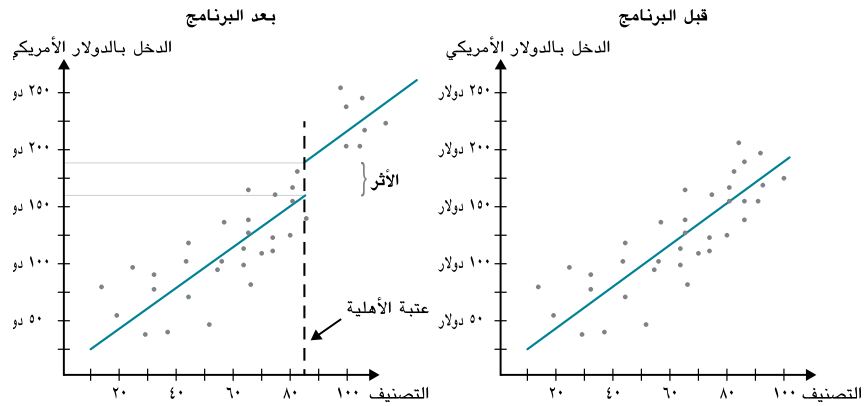
على مجموعة معايير وتخصيص درجة ١ - ١٠٠. إذا أحرز الشباب درجات عند الحد الأدنى أو أعلى من ذلك، مثل ٨٥ فما فوق فإنهم يحصلون على تمويل للبدء أما إذا أحرزوا درجات أدنى فلن يتم قبولهم في البرنامج. تصنيفات الأهلية كذلك يمكن استخدامها لتقييم الأثر.

كيفية العمل

تتمثل فرضية تصميمات تقييم عدم الاستمرارية (أو مؤشر الأهلية) في أن الأشخاص الذين أحرزوا درجات أعلى بالكاد أو أدنى بالكاد من مستوى محدد لا يختلف أحدهم عن الآخر كثيراً، أو على الأقل قد يكون الاختلاف مستمراً عبر النتائج. على سبيل المثال، هل مقدمو الطلبات الذين تلقوا درجة ٨٦ يختلفون كثيراً عما أحرزوا درجة ٨٤؟ على الأرجح لن يختلفوا، أو هل من يبلغون من العمر ١٨ عاماً، الذين قد يكونون مؤهلين لبرامج النقد مقابل العمل يختلفون كثيراً عن أقرانهم ممن يبلغون من العمر ١٧ عاماً، من منهم يكون غير مؤهل؟ إذا كنا في وضع حيث بعض الشباب الذين تلقوا البرنامج (ممن أحرزوا درجات أعلى بالكاد من العتبة) والذين لم يلتحقوا بالبرنامج (أحرزوا درجات أقل بقليل من المستوى المحدد) لم يكونوا مختلفين جوهرياً عن بعضهم البعض ستمكن عندئذ من مقارنة نتائج هاتين المجموعتين، لتحليل أثر البرنامج.

شكل ٦-٦ يوضح ما قد نجده عند تحليل أثر مبادرة القروض الصغيرة للشباب. يشير الرسم البياني على اليسار، في وقت التقدم بطلب للبرنامج، الذين حققوا درجات أفضل كانت بالفعل دخولهم أعلى. قد يكون هناك العديد من الأسباب لذلك، مثل أن من هم أفضل بشكل ما من ناحية التعليم يكسبون أكثر بالفعل وأن تعليمهم أيضاً ساعدهم على ضمان إحراز درجات أفضل، أو أن من لديهم حافز أكبر لبدء الأعمال كانوا بالفعل أكثر تنظيماً للمشاريع، مما يظهر في الدخول الأكثر ارتفاعاً، وأن الحافز أيضاً ساعدهم في إقناع فريق الاختيار بدعمهم. توجد تفسيرات أخرى عديدة ممكنة والتي لا نحتاج بالضرورة لفهمها حتى نطبق هذا الأسلوب.

شكل ٦-٦ نموذج رسم بياني لتصميم عدم الاستمرارية



عند بدء البرنامج، قرر بنك القروض الصغيرة المحلي أن مستوى الحد لتلقي قرض هي ٨٥، وتم قبول أو منع الدعم لجميع المتقدمين بطلبات طبقاً لذلك. نحن نريد الآن أن نعرف إذا ما كان لبرنامج القروض الصغيرة أي أثر على الدخول. كما هو موضح في شكل ٦-٦ (الرسم البياني على الجهة اليمنى)، نفترض أن من أحرزوا درجات أدنى من ٨٥ لديهم نفس النتائج كما في السابق في حين أن دخل من أحرزوا درجة ٨٥ فما فوق زاد عبر اللوحة. من هذه المعلومات، من الممكن تحديد أثر البرنامج، والذي سيتمثل في الاختلاف في النتائج (أي عدم الاستمرارية في العلاقة الخطية) بالقرب من نقطة التوقف.

متى يمكن استخدام تصميم عدم الاستمرارية؟

يمكن استخدام تصميم عدم الاستمرارية بالنسبة لكل من التقييمات المتوقعة أو التي تتم بأثر رجعي. بمعنى آخر، فإنه على عكس الأساليب العشوائية التي تم مناقشتها أعلاه، يمكن أيضاً استخدام هذا التصميم عندما يكون البرنامج في طور التنفيذ أو قد تم إنجازه بالفعل. ويتمثل الشرط الرئيسي لتنفيذ هذا الأسلوب في أن يتم تحديد المشاركة في البرنامج بناءً على قاعدة استهداف محددة بشكل صريح، أي بناءً على مقياس أو علامة دائمة. ولغرض تطبيق هذا الأسلوب، نحتاج إلى الكثير من عمليات المراقبة في المنطقة أعلى وأدنى نقطة التوقف مباشرة، وذلك للحصول على أعداد كافية من الشباب يمكن المقارنة بين بعضهم البعض. وما لم يتم إجراء التقييم بدون بيانات أولية أو ما لم يستغل التقييم سجلات البرامج الكائنة، يتطلب تصميم عدم الاستمرارية عملية جمع للبيانات مشابهة لتلك التي تجرى في تصميم اليانصيب، وبناءً عليه يتطلب ذلك التصميم تكلفة مماثلة.

المزايا

- يستفيد أسلوب عدم الاستمرارية من قواعد الاستهداف الموجودة ولا يشترط أي تغيير في تصميم البرنامج.
- يقدم هذا الأسلوب تقديرات غير متحيزة للمشاركين بالقرب من مستوى الحد.
- لا يتطلب هذا الأسلوب العشوائية من أي نوع. لذا يجوز أن يكون مقبولاً من الناحية السياسية أكثر من الأساليب الأخرى.
- يحدد هذا الأسلوب الآثار المحتملة للقياس الحدي. فعلى سبيل المثال، إذا رأى البرنامج تخفيض عتبة الأهلية من درجة، مثلاً، ٨٥ إلى ٧٥، يمكن لتقييم عدم الاستمرارية أن يوضح ما أثر ذلك على المشاركين، مع تقديم معلومات لتحليل المنافع والتكاليف بالنسبة لهذا الاقتراح.

العيوب

- يتطلب أسلوب عدم الاستمرارية مستوى محدد بدقة لتحديد المجموعات.
- تكون تقديرات الأثر صادقة فقط بالنسبة للهامش بالقرب من نقطة التوقف ولا يمكن تعميمها على الأفراد الذين حصلوا على درجات بعيدة تماماً من مستوى الحد. فلا يقدم الأسلوب أثر متوسط للمشاركين في البرنامج.
- يتطلب هذا الأسلوب عينات تقييم كبيرة حيث إن عمليات الملاحظة حول نقطة التوقف هي التي يمكن استخدامها فقط.
- على النحو الذي تم مناقشته في دوفلو وغلينستر وكريم (٢٠٠٦)، في البلدان النامية، نادراً ما يتم تنفيذ قواعد الأهلية بشكل صارم في المقام الأول، وبناءً عليه، ترتفع فرصة وجود مجموعات غير متميزة، مما يجعل من الصعب الحصول على بيانات صادقة باستخدام هذا الأسلوب.
- بإيجاز، يعتبر أسلوب عدم الاستمرارية حالاً جيداً عندما يبدأ التقييم في مرحلة متأخرة أو عندما يستحيل استخدام الأسلوب العشوائي. ومع ذلك، لا يمكن تطبيقه إلا في حالات محددة. ويعرض مربع ٦-٦ مثالاً على تصميم عدم الاستمرارية.

يستخدم كولينغر وشونديلين (٢٠٠٧) تصميم عدم الاستمرارية لدراسة دور برامج التدريب الخاصة بتنظيم المشروعات على تأسيس الشركات ونواحيها في سياق منافسات خطط الأعمال التي تديرها المنظمة غير الحكومية: تكنوسيرف TechnoServe في أمريكا الوسطى. ويوفر البرنامج خدمات التدريب وتطوير الأعمال وذلك لمساعدة المشاركين في إعداد خطة عمل، وهو يمول عدد مختار من أفضل الخطط.

يستفيد المقيمون من حقيقة أنه للالتحاق بالبرنامج، يشترط أولاً إجراء عملية فرز أولية، تحدد للمتقدمين درجة تصف قدرتهم المحتملة على تنظيم المشروعات. فعدد المتقدمين الذين التحقوا بالبرنامج ثابت قبل بدء المنافسة. ويتم قبول المتقدمين في ورشة العمل إذا كانت درجتهم أعلى من مستوى حد، وإذا كانت الدرجة أقل، يتم رفضهم. وهذا يسمح بمقارنة المستفيدين الذين حصلوا للتو على درجة نجاح بهؤلاء الذين فشلوا في الالتحاق بالبرنامج بهامش صغير. وحيث إن كلا المجموعتين لديهما درجات مماثلة سواء أعلى أو أقل من مستوى الحد، فمن العدالة أن نفترض أنهما تشتركان أيضاً في سمات مشابهة غير ملحوظة، والتي تسمح بدورها بواقع مضاد عالي الجودة.

أكد التحليل الإحصائي أنه قد تم الالتزام بقواعد الأهلية - أي أن الأفراد تم اختيارهم بشكل سليم بناء على درجاتهم - وأن سمات نتائج المتقدمين كانت مستمرة بالإضافة إلى درجاتهم قبل البرنامج. وبعد البرنامج، وجد المقيمون تغييراً أكثر وضوحاً في النتائج حول مستوى الحد. واستناداً إلى تصميم عدم الاستمرارية، استطاعوا أن يثبتوا بدورهم أن التدريب زاد من احتمال إقامة مشروع تجاري بحوالي ١٠ بالمائة ومن احتمال التوسع في مشروع تجاري بأكثر من ٢٠ بالمائة.

أسلوب ٥: الاختلاف في الاختلاف

في الكثير من البرامج، لا يتم إتباع معايير واضحة في اختيار المناطق والمستفيدين المستهدفين. وهذا يمكن أن يؤدي إلى استهداف انتقائي للغاية. فعلى سبيل المثال، قد يكون لدينا معرفة سابقة بمجتمع معين، أو إمكانية أفضل للوصول إلى بعض الأماكن دون غيرها، أو معرفة بشركاء قائمين يمتلكون بنية أساسية بالفعل في موضعها الصحيح والتي نود أن نبني عليها. ومع أنه لا يوجد ما يعيب هذا من حيث المبدأ، إلا أن مثل هذه القواعد الذاتية للاستهداف تجعل من الصعب تطوير واقع مضاد جيد. وبالرغم من ذلك، قد يمكننا الحصول على تقدير تقريبي لتأثير البرنامج باستخدام تصميم التقييم الخاص بالاختلاف في الاختلاف.

كيفية العمل

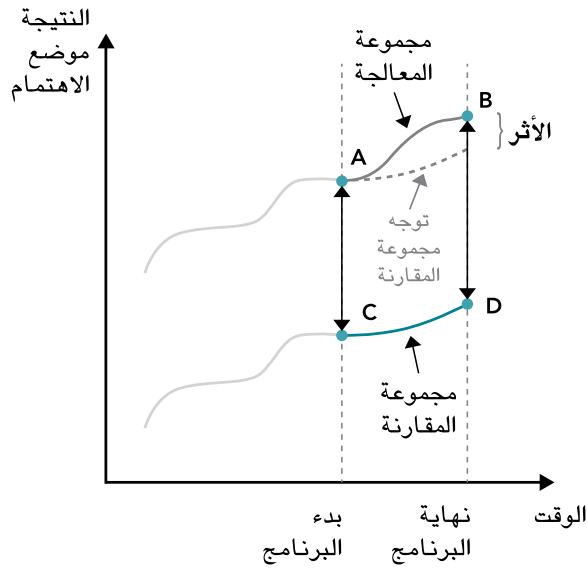
تحديد مجموعة المقارنة. يتكون تصميم الاختلاف في الاختلاف بالأساس مثل «تجربة عشوائية للاختبار القبلي/الاختبار البعدي، ولكنها تفتقر الخاصية الرئيسية لها، ألا وهي التخصيص العشوائي» (تروشم ٢٠٠٦). ففي تصميم الاختلاف في الاختلاف، نحاول تحديد مجموعة مقارنة والتي نرى أنها مشابهة لمجموعة المعالجة المحددة مسبقاً. فعلى سبيل المثال، في التدخلات المركزية لسبل عيش الشباب، قد نختار مركزين تدريب أو فصلين دراسيين متشابهين قابلين للمقارنة. أما في البرامج المجتمعية، قد نستخدم ضاحيتين أو منطقتين متشابهتين. وفي كلا الحالتين، نحاول دوماً اختيار المجموعات التي نعتقد أنها مشابهة لبعضها البعض بقدر الإمكان حتى يمكننا مقارنة المجموعة المعالجة بمجموعة المقارنة بشكل ملائم. ومع ذلك، لا يمكننا أبداً التأكد من أن المجموعتين متشابهتين بشكل دقيق وذلك لأن الاختيار لم يتم بشكل عشوائي - تذكر أن هناك سمات غير ملحوظة لا يمكننا ضبطها - وبناء عليه، يعرف هذا الأسلوب أيضاً باسم تصميم المجموعتين غير المتكافئتين (تروشم ٢٠٠٦).

تقدير الأثر. كما رأينا في بند ٥، فإن المقارنة ببساطة بين نتائج المشاركين وغير المشاركين المختارين بشكل غير موضوعي لا يؤدي إلى أثر البرنامج، حيث إن كل مجموعة مختلفة عن الأخرى في الأغلب. وبالمثل، تمثل مقارنة المشاركين في البرنامج قبل وبعد التدخل إشكالية أيضاً لأن كثير من العوامل الأخرى من المرجح أيضاً أن تؤثر على نتائج المشاركين مع مرور الوقت. ولكن ماذا لو مزجنا الأسلوبين معاً وعقدنا مقارنة قبل وبعد التغييرات التي تحدث في النتائج للمجموعة التي التحقت ببرامجنا وللمجموعة التي لم تشارك؟

دعونا ننصوّر برنامج التدريب في مواقع العمل للشباب. لتطبيق أسلوب تقييم الاختلاف في الاختلاف، يلزم قياس النتائج (على سبيل المثال: الدخل الشهري) بالنسبة لكل من مجموعة المعالجة ومجموعة المقارنة قبل بدء البرنامج (أنظر شكل ٦-٧، نقطتي أ وج) وقياس النتائج لكلا المجموعتين

بعد البرنامج (نقطتي ب و د). وحيث إن كلا المجموعتين من المحتمل أن تكونا مختلفتين منذ البداية، فقد تكون دخولهم عند الأساس مختلفة أيضاً، ولكن ذلك لا يمنع استخدام الأسلوب في الحال. يقارن أسلوب الاختلاف في الاختلاف الفرق في النتائج بين كلا المجموعتين في نهاية التدخل (ب ناقص د) من خلال الفرق في النتائج بين كلا المجموعتين في البداية (أ ناقص ج). ومن جهة أخرى، يمكننا مقارنة الفرق في النتائج للمشاركين (ب ناقص أ) مع خلال الفرق في النتائج لغير المشاركين (د ناقص ج). وينتج عن طرح هذه الفروق من بعضها البعض فكرة تقريبية عن أثر البرنامج، حيث يوضح ما إذا كان برنامج التدريب قد زاد من دخل المشاركين أم لا وكم هذه الزيادة بالنسبة لهؤلاء الذين لم يشاركوا في البرنامج.

شكل ٦-٧ مثال على تحليل الاختلاف في الاختلاف



المصدر: مقتبس من غيرتلير وآخرين (٢٠١١).

افتراض «التوجهات المتساوية». يتمثل الافتراض الضمني لهذا الأسلوب في أنه بالرغم من أن السمات الملحوظة وغير الملحوظة لمجموعتي المعالجة والمقارنة مختلفة بعض الشيء (على النحو المبين في المستويات المختلفة من الدخل في البداية)، فإن اختلافاتهم ثابتة بمرور الوقت أو غير متغيرة زمنياً. وهذا يسمح لنا باستخدام توجه مجموعة المقارنة كتقدير لما كان من الممكن أن يحدث لمجموعة المعالجة الخاصة بنا في غياب التدخل.

هل هذا الافتراض واقعي؟ من المحتمل ألا تتغير الكثير من السمات الملحوظة، مثل سنة الميلاد والنوع وتعليم الأهل والسمات من هذا القبيل، على مدى التقييم. ومع ذلك، لا يمكن تطبيق الافتراض ذاته على عديد من السمات غير الملحوظة مثل السمات الشخصية والدوافع الذاتية للفرد وتفضيلات المخاطر وغيرها، والتي ثبت طبقاً للعديد من الدراسات أنها تتغير بمرور الوقت، ولا سيما فيما يتصل بالبرامج الإنمائية (أنظر، على سبيل المثال، روينز وآخرين ٢٠٠١، وروبرتس وكاسبي، وموفيت ٢٠٠٣). وبناء عليه، لا يمكننا أبدأ التأكد من أن الاختلافات والفروق بين المجموعات لا تتغير بمرور الوقت، والذي بدوره، قد يجعل تقديرات الأثر التي أجريتها متحيزة. حتى وإن بقيت الاختلافات في سمات المشاركين ثابتة، فقد تؤدي هذه الاختلافات إلى تأثيرات تفاعلية مع مرور الوقت. وإذا كان المشاركون في البرنامج، في المتوسط، أكثر تحفظاً وحساساً من غير المشاركين فيه، فيمكن أن يستفيدوا من البرنامج وأن يحققوا بدورهم منافع أعلى من خلال مشاركتهم أكثر مما كان من الممكن لغير المشاركين أن يحققوه. وبالإضافة إلى ذلك، قد تؤثر العوامل الخارجية على المجموعتين بدرجة مختلفة خلال فترة التنفيذ. هكذا سيكون الحال على سبيل المثال إذا بدأت البلدية برنامجاً جديداً في مجتمع المعالجة الخاص بنا، ولكن ليس في مجتمع المقارنة.

[معلومة إرشادية]

هناك اختبار جيد لمعرفة ما إذا كان من الواقعي افتراض التوجهات المتساوية بين المشاركين وغير المشاركين من عدمه، يتمثل في مقارنة التغيرات التي أحدثوها في النتائج قبل تنفيذ البرنامج. فإذا تحركت النتائج تراكباً قبل بدء البرنامج، يمكننا أن نكون أكثر ثقة أن النتائج الخاصة بهم سوف تستمر في هذا الاتجاه أثناء البرنامج. وإذا كانت التوجهات السابقة على البرنامج مختلفة، فقد لا يكون افتراض التوجهات المتساوية صحيحاً. ومع ذلك، فإن معرفة الاختلاف في التوجهات سوف يسمح لنا بضبط هذا الاختلاف على الأقل عند حساب التحليل.

المصدر: مقتبس من غيرتلير وآخرين (٢٠١١).

متى يمكن استخدام تصميم الاختلاف في الاختلاف؟

ويستخدم هذا التصميم بشكل أفضل في غياب آلية استهداف واضحة (مثل التوزيع العشوائي أو ترتيب الأهلية). وحيث إن التصميم يفترض ثبات الاختلافات بين المشاركين وغير المشاركين مع مرور الوقت، يستخدم هذا الأسلوب بشكل أكثر منطقية عندما تكون هناك بيانات جيدة في فترات متعددة قبل بدء البرنامج. ويجب أن يكون هناك ثلاث عمليات لجمع البيانات على الأقل: عمليتين قبل المعالجة، وواحدة على الأقل في النهاية. وهذا يعني أنه لو لم تكن البيانات المتصلة بالمشاركين وغير المشاركين متاحة من خلال قنوات أخرى، مثل مسح الأسر، يمكن أن تكون تكاليف التقييم من هذا القبيل أعلى بكثير من أساليب تقييم الأثر الأخرى.

المزايا

- يوفر تصميم الاختلاف في الاختلاف طريقة لأخذ الاختلافات بين المشاركين وغير المشاركين في الاعتبار.
- يضبط هذا التصميم العديد من التأثيرات الفردية.
- لا يتطلب التصميم تقييماً محتملاً إذا تم بالفعل جمع البيانات اللازمة.
- يعتبر التصميم مفيداً عند استخدامه مع أساليب أخرى لزيادة القوة الإحصائية.

العيوب

- ينتج عن هذا التصميم نتائج أقل موثوقية من أساليب الاختيار العشوائي.
- لا يمكن استخدام هذا التصميم وحده دون افتراض أن مجموعتي المعالجة والمقارنة تتغيران بمرور الوقت بنفس الطريقة.
- يتطلب هذا التصميم ثلاث عمليات لجمع البيانات على الأقل، بينما تتطلب الأساليب الأخرى عمليتين فقط. لذا، فإن هذا التصميم يمكن أن يكون أكثر تكلفة.
- أنظر مربع ٦-٧ للحصول على مثال على هذا التصميم.

عمل كل من ألميدا وغلأسو (٢٠٠٨) على دراسة الآثار قصيرة المدى لبرنامج خاص بتشجيع العمل للحساب الخاص بين المستفيدين من البرامج الإلزامية في الأرجنتين. في أعقاب الأزمة الاقتصادية الشديدة التي حدثت في عام ٢٠٠١، قامت الحكومة الأرجنتينية بتقديم برنامج إلزامي كبير، يدعى جيفيس (Jefes)، يتضمن مبادرة برنامج لتعزيز العمل للحساب الخاص تحت عنوان «المشاريع الإنتاجية الصغيرة». وفر برنامج المشاريع الصغيرة هبات عينية لتمويل المدخلات المالية والمعدات، وكذلك المساعدة الفنية من خلال الزيارات الدورية للمدرسين.

ومن أجل تقييم البرنامج في ظل غياب البيانات التجريبية، قام كل من ألميدا وغلأسو باستخدام إطار عمل الاختلاف في الاختلاف. عمل هذا النهج على مقارنة نتائج سوق العمل بالنسبة للمشاركين في البرنامج قبل وبعد التدخل مع غير المشاركين. ومن أجل تحديد مجموعة مقارنة جيدة، استفاد كل من ألميدا وغلأسو من الحملة الدعائية للبرنامج، والتي تمكن خلالها المستفيدون برنامج جيفيس من التسجيل للتعبير عن اهتمامهم بالبرنامج. من خلال حصر مجموعة المقارنة على الأشخاص الذين أبدوا اهتماماً بمبادرات المشاريع الصغيرة (ولكنهم لم يشاركوا في النهاية)، هدف المؤلفان إلى التقليل من مشاكل مقارنة الأفراد المهتمين بالعمل للحساب الخاص (على سبيل المثال، بسبب قدرتهم على تنظيم المشاريع أو وجود الحافز لديهم) بالأفراد غير المهتمين.

تم تنفيذ مسح أسر أولي مع ٣٠٩ مشارك و٢٤٤ شخص من غير المشاركين في تشرين الثاني/ نوفمبر ٢٠٠٤. تمت إدارة المسح من قبل سيمبرو (SIEMPRO)، الوكالة الأرجنتينية العامة لمتابعة وتقييم برامج الفقر. تمت مقابلة نفس الأسر مرة أخرى بعد مرور عام واحد وكان ذلك في نهاية عام ٢٠٠٥. في ظل وجود فقط مجموعتي بيانات، كان على المقيمين أن يفترضوا أنه في ظل غياب البرنامج، كان لا بد أن يكون لدى كل من المشاركين وغير المشاركين توجهات قابلة للمقارنة في نتائج سوق العمل (افتراضات «التوجهات المتساوية»).

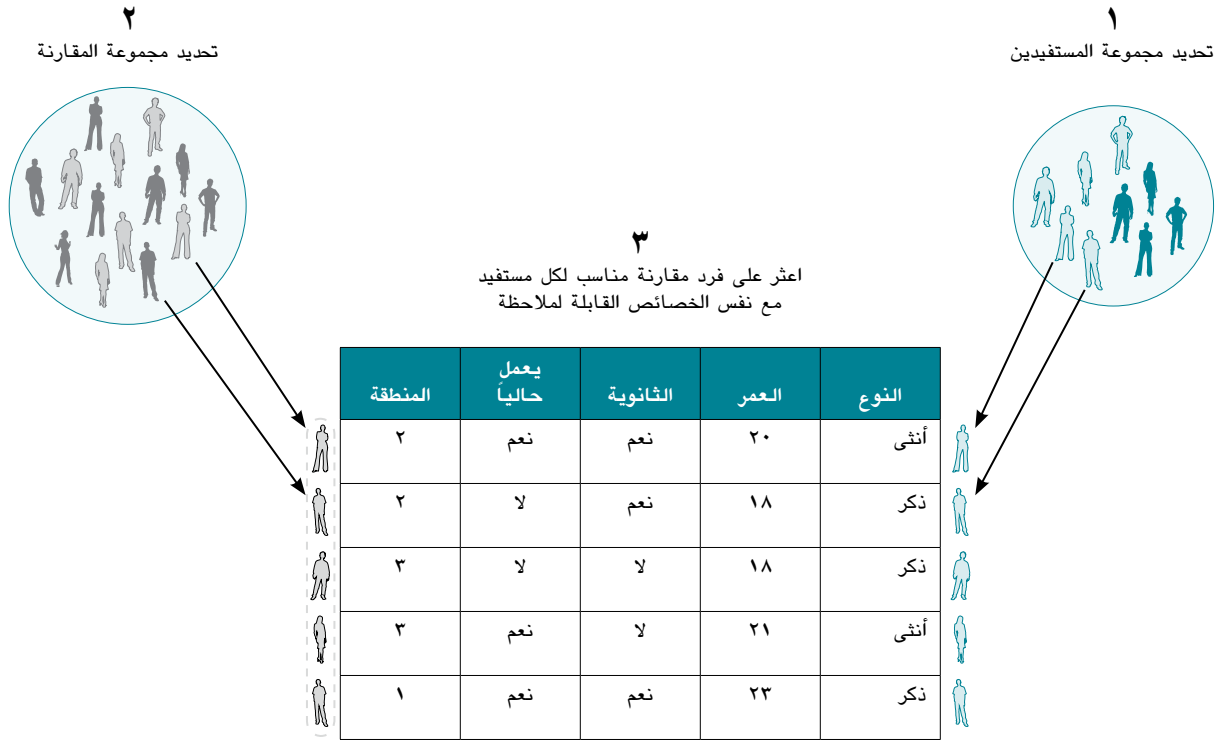
أشارت النتائج إلى أنه نظراً لمعدل المشاركة المنخفض نسبياً، فإن تحريك العمل للحساب الخاص من خلال توفير رأس المال الإبتدائي والتدريب على الأعمال لا يكون بالضرورة خياراً جذاباً بالنسبة لجميع المستفيدين من البرامج الإلزامية. وفضلاً عن ذلك، على الرغم من أن البرنامج أدى إلى زيادة عدد ساعات العمل للمشاركين، إلا أنه أخفق في زيادة متوسط الدخل بشكل كبير. وأخيراً، لم يستفد كل شخص من البرنامج بنفس القدر، وتمت ملاحظة وقياس الآثار الإيجابية فقط على المشاركين الحاصلين على مستويات أعلى من التعليم.

أسلوب ٦: المطابقة

كما هو بالنسبة لتصميم أسلوب الاختلاف في الاختلاف، يتم استخدام المطابقة في حالة عدم وجود قواعد صارمة تقسيم أخرى محددة من قبل البرنامج. في الماضي، كان أسلوب المطابقة متعارف عليه بين المختصين في مجال تقييم البرامج، ولكن الآن أصبحت هناك أساليب أكثر حيثية، مثل الأساليب المبنية أعلاه، تفوق على هذا الأسلوب.

كيفية العمل

يعمل أسلوب المطابقة على عقد مزاجية بين الشباب المشاركين في برنامج ما وبين غير المشاركين بناءً على الخصائص التي يمكن ملاحظتها (العمر، الجنس، مستوى التعليم، الوضع الوظيفي والسكن وغيرها من العوامل). وهذا يعني أنه لكل شاب (أو مجموعة من الشباب) في مجموعة المعالجة، يقوم المطابقة بعمل وحدة مقارنة مصطنعة يكون فيها أكبر عدد ممكن من الصفات المتشابهة (أنظر شكل ٦-٨). تحاول هذه التقنية الإحصائية محاكاة مجموعة مقارنة لا وجود لها بغير هذه الكيفية. في النهاية، يمكن مقارنة متوسط النتائج للحاصلين على المعالجة مع نتائج مجموعة المقارنة، وينشأ أثر التدخل عن الاختلاف بينهما.



يتطلب تحديد نظير جيد لكل مشارك في البرنامج إيجاد الخصائص التي تفسر قرار فرد ما التسجيل في البرنامج. وللأسف، فهذا الأمر ليس سهلاً كما يبدو. وكما يبين غير تالر وزملائه (٢٠١١)، إذا كانت قائمة الخصائص ذات الصلة صغيرة (كما في شكل ٦-٨ أعلاه)، فمن المحتمل أن نجد نظيراً لكل شاب من مجموعة المعالجة، ولكن قد لا يكون كل نظير دقيق بشكل بارز وتعرض لخطر ترك أية معايير أخرى قد تكون مهمة. وإذا أردنا في المقابل المطابقة بناء على عدد كبير من الخصائص (بإضافة، على سبيل المثال، المستوى التعليمي للأهل ودرجات الاختبار ومستوى الدخل)، فقد يكون من الصعب تحديد نظير لكل وحدة من الوحدات في مجموعة المعالجة إلا إذا كان عدد الملاحظات في قاعدة البيانات الخاصة بشباب المقارنة كبيراً جداً.

متى يمكنني استخدام أسلوب المطابقة؟

يمكن استخدام أساليب المطابقة في العديد من الحالات، بغض النظر عن التغطية التي يوفرها البرنامج أو معايير الاستهداف. عادة ما تستخدم هذه التقنيات عملياً عندما لا يكون أي من تصاميم التقييم مجدياً، لا سيما عندما يبدأ التقييم بعد التنفيذ. ونظراً لعدم قدرة أسلوب المطابقة على السيطرة على الخصائص التي لا يمكن ملاحظتها، إلا أنه يفضل استخدام هذا الأسلوب مع أحد أساليب التقييم الأخرى. وكذلك، كي تكون المطابقة صحيحة نحتاج عادة إلى حجم عينة كبير لضمان إيجاد مجموعة مقارنة قابلة للمطابقة (أنظر مربع ٦-٨). إذا لم يتم جمع البيانات المطلوبة من خلال قنوات أخرى، فقد يكون التقييم مكلفاً أكثر من الطرق الأخرى المبنية في هذا البند.

[معلومة إرشادية]

يمكن التغلب على التحدي المتعلق بإيجاد أزواج في مجموعات المعالجة والمقارنة لديها العديد من الخصائص التي يمكن مقارنتها، من خلال استخدام أسلوب يعرف باسم درجة مطابقة الميل. بدلاً من مطابقة وحدات المعالجة والمقارنة بناء على نفس الخصائص لجميع المعايير التي تم اختيارها، يحسب أسلوب درجة مطابقة الميل احتمال (درجة ميل) تسجيل كل شاب في البرنامج بناء على العديد من الخصائص التي تمت ملاحظتها. وبمجرد حساب درجة الميل (عدد يكون بين ٠ و ١) لجميع المشاركين وغير المشاركين الذين تتوفر بياناتهم، يتم مطابقة المشاركين مع غير المشاركين الذين لديهم الدرجة الأقرب. عندئذ، يشكل هؤلاء الأشخاص المتوافقين من غير المشاركين في البرنامج، مجموعة المقارنة.

مربع ٦-٨ خطوات تطبيق أسلوب المطابقة

١. حدد الشباب الذين سجلوا في البرنامج والذين لم يسجلوا.
٢. اجمع معلومات متعمقة حول الخصائص القابلة لملاحظتها (مثل العمر والنوع ومستوى التعليم والوضع الوظيفي) لكل من الشباب المسجلين وغير المسجلين من خلال مسح أساسي أو من خلال الرجوع إلى البيانات الموجودة.
٣. باستخدام أسلوب مطابقة إحصائي مثل مطابقة درجة الميل، قم بإيجاد نظير لكل مشارك بشخص من غير المشاركين لديه خصائص متشابهة.
٤. قارن نتائج الشباب المسجلين ونتائج المقارنات المتطابقة معهم. يكون الاختلاف في النتائج هو أثر البرنامج على ذلك الفرد بعينه.
٥. احسب المتوسط التقديري لأثر البرنامج من خلال أخذ متوسط الآثار الفردية.

المزايا

- تسمح المطابقة بمقارنة النتائج بين الأفراد المتشابهين.
- يمكن استخدامها مع أساليب أخرى للتحقق من جودة مجموعة المقارنة.

العيوب

- لأن المطابقة تتطلب عقد مقارنات مباشرة بين الأشخاص، فقد تكون هناك حاجة إلى إجراء مسح واسع للعينة لاشتقاق مجموعة مقارنة مناسبة.
 - يمكن إجراء المطابقة على الخصائص التي يمكن ملاحظتها فقط. لا يمكن دمج السمات التي لا يمكن ملاحظتها مثل الشخصية والحافز والدعم العائلي وغير ذلك في هذا الأسلوب. وبالتالي فإن المطابقة تتطلب افتراض عدم وجود اختلافات منهجية في الخصائص التي لا يمكن ملاحظتها بين مجموعتي المعالجة والمقارنة، وعادة ما يكون هذا الأمر غير قابل للتصديق. وإذا لم يتم دعم هذا الفرض، فقد تؤدي المطابقة إلى التحيز في تقدير أثر البرنامج.
 - قد لا يكون ممكناً العثور على توافق مناسب لكل فرد في مجموعة المعالجة، مما يضعف الصدق الخارجي لتقدير الأثر.
- للحصول على مثال على المطابقة، أنظر مربع ٦-٩.

استخدم كل من غاراميلو وبارودي (٢٠٠٣) مطابقة درجة الميل لتقييم برنامج تنظيم الأعمال الشبابية الذي نفذته منظمة كوليكثيفو انتيغرا دي ديسارولو، منظمة غير حكومية في البيرو. ولتقدير أثر مسابقة خطة العمل وخدمات الدعم التابعة التي تتكون من التدريب والمتابعة وفرص التدريب العملي للمشاركين، كون المقيمون مجموعة مقارنة تتكون من الشباب الذين شاركوا في النشاطات التحضيرية للبرنامج (ما قبل التدريب) ولكنهم إما لم ينضموا لمسابقة خطة العمل أو لم يقدموا مشروعات مقترحة فائزة.

قام المقيمون بحساب احتمال أن تكون مشاركة شخص ما في البرنامج قائمة على أساس الخصائص القابلة لملاحظتها مثل العمر والنوع والمستوى التعليمي والحالة الاجتماعية. وبعد ذلك تمت مطابقة كل مستفيد من شخص من مجموعة المقارنة له درجة ميول شبيهة. تم بعد ذلك استخدام مقارنة النتائج (من حيث استدامة العمل وعدد فرص العمل التي تم إيجادها والدخل) بين المستفيدين وأقرانهم المتوافقين لتقدير أثر التدخل.

ومع ذلك، وحيث إن المطابقة لا تقوم إلا على الخصائص القابلة لملاحظتها، كانت هناك فرصة حقيقية لأن تكون الآثار الإيجابية التي تم تحديدها في التقييم مبالغة للأثر الحقيقي للتدخل. ففي الواقع، من المرجح أن يكون الشباب الذين شاركوا بنجاح في مسابقة خطة العمل مختلفين عن الشباب في مجموعة المقارنة من حيث، على سبيل المثال، تحفزهم أو مستوى مهاراتهم، وقد يكونون منظمين أكثر نجاحاً من أقرانهم حتى وإن لم يشاركوا في برنامج تنظيم الأعمال.

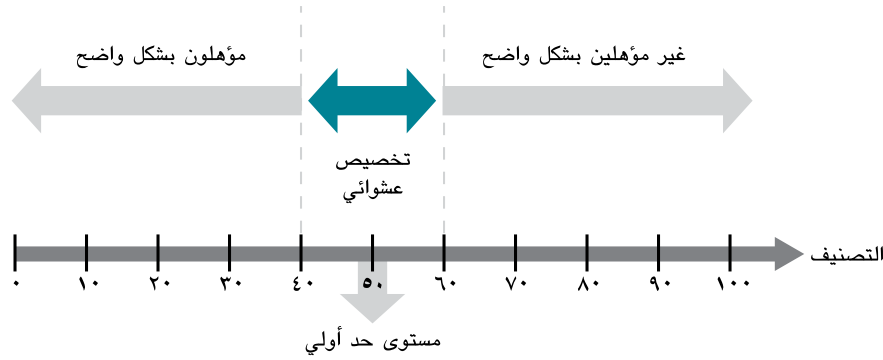
الجمع بين الأساليب

كما رأينا، بعض الأساليب تكون أقوى في إنشاء واقع مضاد أكثر من غيرها. على وجه الخصوص، قد يكون من الصعب إيجاد مجموعة مقارنة جيدة عندما لا يكون قد تم التخطيط للتقييم منذ بداية البرنامج. إن الجمع بين الأساليب قد يعوض بعض نقاط الضعف في أحد الأساليب ويزيد من صدق الواقع المضاد الذي تم تقديره.

تصميم الانقطاع العشوائي

يجمع هذا الأسلوب بين تصميم الانقطاع والتخصيص العشوائي. إذا لم يتم تحديد الانقطاع بوضوح، أو إذا لم يتم تبريره بشكل كاف، فمن الممكن أن يكون الترتيب عشوائياً حول الانقطاع. وفي هذه الحالة، يستمر الشباب الذين يكونون مؤهلين بشكل واضح في الاستفادة من البرنامج بينما لا تقدم خدمات البرنامج للشباب غير المؤهلين بشكل واضح للاستفادة من البرنامج (أنظر شكل ٦-٩). يتم فقط اختيار المجموعة التي تكون قريبة من مستوى الحد، لعملية الترتيب العشوائي. بهذه الطريقة، قد يحصل بعض الأشخاص الذين قد لا يكونوا استفادوا من البرنامج بطريقة مغايرة، على خدمات البرنامج الآن، والعكس. كما هو الحال في تصميم الانقطاع، تكون النتائج صحيحة فقط بالنسبة لهؤلاء المشاركين فقط عند هامش القبول. ومع ذلك، نظراً للترتيب العشوائي المعين، يمكننا أن نكون أكثر ثقة من أن مجموعتي المعالجة والمقارنة تشتركان في الخصائص ذات الصلة. ونحن نحتاج إلى حجم عينة أصغر للحصول على نتائج ذات دلالة إحصائية. يتم بعد ذلك إجراء التحليل بنفس طريقة إجراء أي تصميم عشوائي. حيث تتم مقارنة متوسط نتيجة الأفراد في مجموعة المعالجة مع متوسط نتيجة الأفراد في مجموعة المقارنة، ويكون الاختلاف هو الأثر السببي للبرنامج على الأفراد الذين تم اختيارهم.

شكل ٦-٩ نطاق الأهلية (مثال على ترتيب درجة الفقر



ملاحظة: الدرجة الأقل تمثل مستوى أعلى من الفقر.

الاختلاف في الاختلاف أو المطابقة المصحوب بالعشوائية

يفترض أسلوب الاختلاف في الاختلاف أن الأفراد في مجموعتي المعالجة والمقارنة متشابهين إلى حد كبير، أو على أقل تقدير الاختلافات بينهم ثابتة بمرور الوقت. وعلى نحو مماثل، يفترض أسلوب المطابقة وجود خصائص مماثلة يمكن ملاحظتها مما يبرر إجراء مقارنة بين فردين. بينما لا تتطلب العشوائية تطبيق أي من الافتراضين من أجل تقدير أثر البرنامج. إلا أنه يمكن تحسين العشوائية عند استخدامها مقترنة مع أي من الأسلوبين أو كلاهما. ومع الحد من الاختلافات بين الأفراد الذين يتم إجراء المقارنة بينهم، يعمل كل من أسلوب الاختلاف في الاختلاف وأسلوب المطابقة على زيادة القوة الإحصائية دون الحاجة إلى زيادة عدد المشاركين. ومن خلال الجمع بين الأساليب غير العشوائية بمثلتها العشوائية، يمكن خفض تكاليف المسح.

الجمع بين أسلوب الاختلاف في الاختلاف والمطابقة

إذا لم يتم تنفيذ أي نوع من العشوائية أو تصميم عدم الاستمرارية، فثمة طريقة أخرى ألا وهي الجمع بين أسلوب الاختلاف في الاختلاف والمطابقة ومن ثم تخفيف بعض نقاط الضعف في كلا الأسلوبين عند استخدام كل منهما على حدة. نظراً لأن أسلوب الاختلاف في الاختلاف لا يمكنه أن يضمن التكافؤ بين مجموعتي المعالجة والمقارنة، فإن الجمع بين هذا الأسلوب وبين عملية مطابقة بسيطة أو مطابقة درجة الميل يمكنه على الأقل ضمان أن كلتا المجموعتين متشابهتين إلى حد كبير فيما يتعلق بالخصائص القابلة للملاحظة.

للحصول على نظرة عامة على أساليب التقييم القياسية، أنظر جدول ٦-١.

[معلومة إرشادية]

من الناحية العملية، يتعين على كبير المقيمين تقييم ما إذا كان الجمع بين الأسلوبين سيعود بالنفع أم لا. ومن ثم لا يحتاج الممارسون إلى الشعور بالقلق بشأن تفاصيل المناهج المرتبطة بل عليهم الإدراك الواعي أن هذا الجمع قد يكون طريقة للحصول على تقديرات للأثر أكثر موثوقية.

جدول ٦-١ نظرة عامة على أساليب تقييم الأثر

الأسلوب	الوصف	مجموعة المقارنة	الاقتراحات المعملوية	البيانات اللازمة	متى يمكن استخدامه
تصميم اليانصيب	يتم تخصيص عينة من الأشخاص المؤهلين لمجموعة الأفراد الذين يحصلون على التدخل ومجموعة الأفراد الذين لا يحصلون عليه. ويكون الأثر هو الاختلاف في النتائج بين المجموعتين.	الأشخاص المختارون عن طريق اليانصيب.	• تعد العشوائية وسيلة ناجحة ويمكن الالتزام بها. • تكون المجموعتين متطابقتين إحصائيا في عوامل ملحوظة وغير الملحوظة.	• بيانات ما بعد التدخل لمجموعتي المعالجة والمقارنة. • البيانات الأولية مرغوبة..	• إذا أمكن تصميم الدراسة قبل بدء البرنامج. • إذا كانت الموارد المتاحة غير كافية، ومن الأهمية بمكان ضمان وجود أساليب عادلة من أجل تسجيل ذوي الحاجة بالبرنامج. • في حالة استحالة التحاق مجموعة المقارنة بالبرنامج نتيجة طول التقييم.
التصميم العشوائي المنفذ على مراحل	يخصص الأشخاص المؤهلين في شرائح للمعالجة للحصول على البرنامج على نحو متعاقب.	المرجعين بقائمة الانتظار.	• تعد العشوائية وسيلة ناجحة ويمكن الالتزام بها. • لن يقوم الأشخاص المدرجين في مراحل البرنامج في وقت لاحق بتغيير سلوكهم بشكل ملحوظ أثناء انتظار المشاركة.	• بيانات ما بعد التدخل لمجموعتي المعالجة والمقارنة. • البيانات الأولية مرغوبة.	• إذا أمكن تصميم الدراسة قبل بدء البرنامج. • إذا كانت الموارد المتاحة غير كافية، ومن الأهمية بمكان ضمان وجود أساليب عادلة من أجل تسجيل ذوي الحاجة بالبرنامج. • إذا كان البرنامج يتم تقديمه على فترات زمنية.
التصميم العشوائي الترويجي	يتم تشجيع مجموعة عشوائية من الأفراد أو المجموعات على التسجيل بالبرنامج. ويتم قياس الأثر عن طريق مقارنة متوسط نتائج الأفراد الذين تم تشجيعهم بغيرهم ممن لم يتم تشجيعهم.	الأفراد الذين لم يحصلوا على التشجيع.	• تعد العشوائية وسيلة ناجحة ويمكن الالتزام بها. • سيكون هناك تقلبات في طريقة التزام من تم تشجيعهم ومن لم يتم تشجيعهم.	• بيانات ما بعد التدخل لمجموعتي المعالجة والمقارنة. • البيانات الأولية مرغوبة.	• إذا أمكن تصميم الدراسة قبل بدء البرنامج. • إذا لم يمكن استبعاد أحد من المشاركة في البرنامج. • إذا كانت المشاركة تطوعية، ومن المتوقع مشاركة المزيد من المشاركين في البرنامج إذا تم تشجيعهم على الالتحاق بالبرنامج.

جدول ٦-١ نظرة عامة على أساليب تقييم الأثر

الأسلوب	الوصف	مجموعة المقارنة	الإفتراسات المطلوبة	البيانات اللازمة	متى يمكن استخدامه
تصميم عدم الاستمرارية	يتم تصنيف الأفراد وفقاً لمعايير محددة وقابلة للقياس. هناك حد فاصل يحدد الأفراد المؤهلة للمشاركة. يتم عندئذ مقارنة نتائج المشاركين وغير المشاركين القريبة من مستوى الحد هذا، ومن ثم يتم ضبط معيار الأهلية.	الأفراد القريبين من مستوى الحد ممن لم تتسن لهم المشاركة.	يشترك الأفراد القريبون من مستوى الحد على جانبيه في الخصائص الملحوظة وغير الملحوظة.	• بيانات ما بعد التدخل للأفراد الأقرب للحد الفاصل. • البيانات الأولية مرغوبة.	• إذا لم يكن استخدام الأسلوب العشوائي ممكناً، أو إذا بدأ التقييم بعد بداية البرنامج. • إذا كان الاختيار بناءً على التصنيف المستمر بجانب مستوى الحد.
الاختلاف في الاختلاف	تتم مقارنة نتائج المشاركين في التدخل البرنامج وغير المشاركين به قبل التدخل وبعده، ويكون التغير النسبي في النتائج هو أثر البرنامج.	مجموعة غير متكافئة من الأفراد الذين لم يشاركون في البرنامج، ولكن تم جمع البيانات الخاصة بهم.	يمرور الوقت، نلاحظ أن أفراد مجموعة المعالجة لا يتغيرون تغيراً جذرياً بطريقة تختلف عن أفراد مجموعة المقارنة.	• البيانات الأولية وبيانات المتابعة. • أفضل الحصول على البيانات من ثلاث فترات زمنية كحد أدنى (يجب الحصول على اثنين منهم قبل بدء البرنامج).	• إذا بدأت الدراسة بعد بدء البرنامج. • إذا وجد أفراد من غير المشاركين متشابهين مع المشاركين. • إذا لم تؤخذ عدالة الاختيار في الحسبان. • يحقق أفضل النتائج عند استخدامه مع غيره من الأساليب.
أسلوب المطابقة	يتم المطابقة بين الأشخاص من مجموعة المعالجة مع غيرهم من غير المشاركين ممن لديهم صفات متشابهة قابلة لملاحظتها.	مطابقة تامة لكل مشارك، على الأقل واحد من غير المشاركين سيمثل تطابقاً مع المشارك وفقاً للخصائص المختارة.	يستطيع الباحثون تحديد كافة خصائص الأشخاص ذات الصلة من خلال مسح.	• مسح كبير (تعداد السكان ومسوح القوى العاملة ومسوح الأسر) يمكن الحصول على النتائج المثلّي في حالة تطبيقه مع مسح الأسر الخاصة بالبرنامج (الوضع الأمثل ملاحظتان لكليهما)	• إذا بدأت الدراسة بعد بدء البرنامج. • إذا وجد أفراد من غير المشاركين متشابهين مع المشاركين. • إذا لم تؤخذ عدالة الاختيار في الحسبان. • يحقق أفضل النتائج عند استخدامه مع غيره من الأساليب.

النقاط الرئيسية

١. تسمح مجموعة مختارة فقط من أساليب تقييم الأثر بالحصول على نتائج موثوق بها للواقع المضاد والتي يمكن التعويل عليها.
٢. تصميم اليانصيب والتصميم العشوائي المنفذ على مراحل والتصميم العشوائي الترويجي وكذلك تصميم عدم الاستمرارية جميعها تنتج تقديرات الواقع المضاد وذلك من خلال قواعد تخصيص محددة للبرنامج. ويوفر أسلوب الاختلاف في الاختلاف والمطابقة للمقيم أدوات إضافية - إلا إنها أقل دقة - لتقييم الأثر وذلك عند بدء التقييم بعد التنفيذ وعندما تكون معايير التأهيل محددة بشكل أقل وضوحاً.
٣. لا يوجد أسلوب واحد هو الأفضل لكل برنامج. وإنما يعتمد أفضل أسلوب على السياق التشغيلي (مثل: التوقيت والتغطية والاستهداف) للبرنامج. ومن ثم، يحتاج مدراء البرنامج إلى مناقشة القيود البراجمجة مع متخصص في التقييم وذلك لما سيكون لهذه القيود من أثر على قابلية تنفيذ مختلف تصميم التقييم.
٤. كلما كان ذلك ممكناً، نوصي بشدة بوضع خطة لتقييم الأثر قبل بدء تنفيذ البرنامج. حيث تميل التقييمات بأثر رجعي لأن تكون أقل قوة وقد لا تكون ممكنة على الإطلاق إذا لم يتم جمع البيانات اللازمة من خلال قنوات أخرى.
٥. في بعض الحالات، ربما تكون الأساليب المطروحة هنا غير قابلة للتنفيذ بسبب متطلبات الميزانية أو قيود التوقيت أو بسبب قضايا سياسية.

دراسة حالة لـ (NUSAF): اختيار تصميم اليانصيب

تم العمل على تطوير تقييم أثر برنامج NUSAF أثناء فترة إعداد البرنامج، ولأن عدد المتقدمين المؤهلين إلى البرنامج فاق بكثير سعة تمويل البرنامج، توقف تصميم تقييم الأثر على توافر عدد كبير من المتقدمين المؤهلين لبرنامج فرص الشباب ولكن بدون إمكانية تمويلهم. ومع تلقي العدد الضخم من الاشتراكات في البرنامج، قررت إدارة NUSAF ومنسقي البرنامج أن اختيار المستفيدين عن طريق نظام اليانصيب لم يكن مجد فقط بل أنه كان بمثابة آلية عادلة وشفافة تعمل على توزيع التمويل بين المتقدمين المؤهلين من الشباب بشكل متساوي.

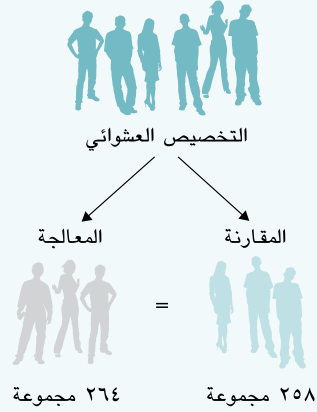
تلقى المسؤولون الفنيون للمنطقة في NUSAF تعليمات للتأكد من أن الطلبات المقدمة تفي بالحد الأدنى من المعايير الفنية المطلوبة للأهلية وكذلك لإجراء تقييمات ميدانية على البرامج التي سيتوقع عليها الاختيار من أجل التمويل. وقد تم إرسال قائمة بالبرامج المؤهلة والمؤكد صحة إلى وحدة إدارة المشروع وتم تقديمها فيما بعد إلى فريق تقييم الأثر، والذي أجري بدوره عملية اليانصيب لاختيار المشروعات المقترحة الممولة. وفي كل مقاطعة، تم اختيار ٣٠-٦٠ بالمائة من المجموعات المؤهلة للحصول على التمويل، وذلك حسب حدود الميزانية الخاصة بتلك المقاطعة تحديداً.

العينة (٥٢٢ مجموعة مؤهلة)

فور استلام قائمة بجميع مقدمي الطلبات من المسؤولين الفنيين للمقاطعة، تم الانتهاء من إجراء تخصيص عشوائي للمتقدمين لمجموعتي المعالجة والمقارنة بشكل كامل لكل مقاطعة. ثم تم تعيين رقم عشوائي لكل متقدم في المجموعة باستخدام محدد أرقام عشوائي. ثم تم تخصيص المجموعات من أول شاب حتى آخر شاب باستخدام الرقم العشوائي. تم حساب مجموع تكاليف البرنامج. بدءاً من المشاريع الأولى التي تم اختيارها عشوائياً، تم منح التمويل إلى المشاريع حتى نفاذ مجموعة الموارد المتوافرة الخاصة بهذه المنطقة. وتبقى المشروعات الأخرى غير ممولة ويتم تخصيصها لمجموعة المقارنة.

أثناء سير العملية، تم اختيار ٢٦٤ مشروع لتمويله وتشتمل هذه المشاريع على مجموعة المعالجة. أما بالنسبة إلى مجموعة المشروعات المتبقية والتي يبلغ عددها ٢٥٨ مشروع مؤهل والتي لم يتم اختيارها للتمويل فقد شكلت مجموعة المقارنة. ومن أجل إجراء تقييم الأثر، أتاح تكوين مجموعة مقارنة متكافئة تقدير الواقع المضاد، وهي الحالة التي كانت سيتم تجربتها على مجموعة المعالجة في غياب المعالجة.

المصدر: باتمان ومارتينيز وفيالا (٢٠٠٩).



قراءات رئيسية

Duflo, E., Glennerster, R. and Kremer, M. 2006. "Using Randomization in Development Economics Research: A Toolkit." BREAD Working Paper No. 136. (For advanced readers.)

<http://www.povertyactionlab.org/sites/default/files/documents/Using%20Randomization%20in%20Development%20Economics.pdf>.

Gertler, P., Martinez, S., Premand, P., Rawlings, L., and Vermeersch, C. 2011. *Impact Evaluation in Practice*. Washington, DC: The World Bank. (Chapters 4–8 are relevant to this note.) <http://www.worldbank.org/ieinpractice>

Khandker, S., Koolwal, G., and Samad, H. 2010. *Handbook on Impact Evaluation: Quantitative Methods and Practices*. Washington, DC: The World Bank. (For advanced readers. Chapters 3–7 are relevant to this note.)
http://www-wds.worldbank.org/external/default/WDSContentServer/IW3P/IB/2009/12/10/000333037_20091210014322/Rendered/PDF/S20990PUB0EPI1101Official0Use0Only1.pdf

بند

بند ٧: دليل خطوة بخطوة لتقييم الأثر

إن هذا البند هو دليل خطوة بخطوة لتنفيذ تقييم للأثر لتدخلات سبل عيش الشباب. لن نحل هذه المعلومات في هذا البند محل اختصاصي تقييم الأثر، حيث سيكون هناك حاجة دائمة له لتنفيذ تقييم مناسب. وبدلاً من ذلك، سيُسَهِّل هذا البند التخطيط لتقييم الأثر من منظور البرنامج، من التحضير إلى نشر نتائج التقييم (أنظر شكل ٧-١). وعلاوة على ذلك فإنه سيُوضَّح أدوار ومسؤوليات أصحاب المصالح المشاركين في التقييم. نحن نأمل في إزالة الغموض عن معنى إجراء تقييم للأثر وبهذا جعل التفكير بإجراء تقييم الأثر أسهل لكل منظمة أو برنامج.

شكل ٧-١ خطوات لإجراء تقييم للأثر



* لم تنطبق هذه الخطوة فقط على الأساليب التي تتطلب من المنظمة جمع بيانات.

لنقاط الرئيسية

توضح البنود ٢-٦ من هذا الدليل الخطوات التي يجب اتخاذها قبل بدء تقييم للأثر. اطرح الأسئلة التالية:

- هل قمت بالتحديد الواضح لهدف برنامجي؟ يُمثل هدف البرنامج ما نريد تحقيقه، والنتيجة المنشودة من تدخلنا. وكلما كان الهدف محدداً من حيث الفئة المستهدفة والحجم وتوقيت التغييرات المتوقعة، سيكون من الأسهل تتبع التقدم وتنفيذ تقييم ما. على سبيل المثال: «بحلول عام ٢٠١٥، مضاعفة دخل ١,٠٠٠ من الشباب غير المتحقين بالدراسة في ليما، بيرو» (انظر بند ٢).
- هل عملت على تحضير سلسلة نتائج؟ تعطي سلسلة النتائج أصحاب المصالح موجزاً منطقياً ومعقولاً حول كيف يمكن لمصادر ونشاطات البرنامج أن تؤدي إلى النتائج المنشودة وتلبي هدف البرنامج. ويجب على كل برنامج إدراج سلسلة النتائج الخاصة به خطياً بما أنها أساس المتابعة، فضلاً عن تحديد أسئلة التقييم (انظر بند ٣).
- هل قمت بوضع نظام متابعة مع مؤشرات وآليات لجمع البيانات؟ يجب أن يوجد في كل تدخل نظام متابعة قبل بدء تقييم للأثر. ويتطلب نظام المتابعة مؤشرات محددة وتقنيات لجمع البيانات في جميع مستويات سلسلة النتائج وذلك من أجل تتبع التنفيذ والنتائج. وبدون وجود نظام متابعة جيد، قد تكون نتائج تقييم الأثر ذات فائدة محدودة لأنه سيكون من المستحيل تحديد ما إذا كانت النتائج غير المرضية المحتملة هي بسبب تصميم البرنامج السيء أو ببساطة بسبب التنفيذ السيء (انظر بند ٣).
- هل قمت بتدوين أهداف التعلم وأسئلة التقييم؟ يجب أن يعتمد تقييم الأثر على احتياجاتنا للمعلومات، حيث نجيب تقييمات الأثر على أسئلة السبب والنتيجة، أي أنها تحدد إذا كانت نتائج البرنامج المحددة (وهي عادة مجموعة فرعية من تلك المحددة في سلسلة النتائج) هي نتيجة التدخل. وبما أن نوع الأسئلة التي نرغب بالإجابة عنها قد يختلف، قد نحتاج إلى التفكير بأدوات تقييم أخرى أبعد من تقييم الأثر من أجل الإجابة على جميع أسئلتنا (انظر بند ٤).
- هل قمت بتحديد مجموعة أساليب لتقييم الأثر؟ قبل البدء، يجب أن يكون لدينا فهماً أساسياً للجانب التقني العام لتقييم الأثر والمنهجيات الرئيسية التي يمكن استخدامها. وعندما نعرف البرنامج الذي سيتم تقييمه، نستطيع تحديد المنهجية الأكثر ملاءمة لسياقنا التشغيلي. وسيساعد هذا الحد الأدنى من الفهم في النقاشات اللاحقة مع خبراء التقييم كما أنه سيسهل من التخطيط (انظر بند ٥ وبند ٦).
- وعملياً، يوجد عادة سوء فهم بين مدراء البرنامج وخبراء تقييم الأثر لأنه لم يتم التعريف الواضح لسياق التقييم مقدماً. إن امتلاك فكرة واضحة حول الكيفية المنشودة لعمل التدخل والأمور التي يجب تعلمها من التقييم ما سيجعل من الخطوات اللاحقة أكثر فعالية وأكثر توفيراً للوقت والمال.

[معلومة إرشادية]

لكي ترى إذا كان برنامجك جاهزاً لتقييم للأثر ولمساعدتك في تحديد أسلوب التقييم المناسب، قد ترغب بالمشاركة في ورشة عمل عن تقييم الأثر يمكنك فيها استشارة الخبراء حول تفاصيل برنامجك. وتوفر المنظمات التالية مثل مراكز المساعدة والمشورة هذه:

شبكة توظيف الشباب

<http://www.ilo.org/public/english/employment/yen/whatwedo/projects/clinics.htm>

معمل عبد اللطيف جميل

للتطبيقات العملية لمكافحة الفقر (J-PAL)
<http://www.povertyactionlab.org/course>

البنك الدولي

<http://web.worldbank.org/WBSITE/EXTERNALTOPICS/EXTPOVERTY/EXTISPMA/O,contentMDK:21754074~menuPK:384336~pagePK:148956~piPK:216618~theSitePK:384329,00.html>

تحديد الجدول الزمني والميزانية

الجدول الزمني

بحسب التعريف، يعتمد توقيت تقييم الأثر اعتماداً كبيراً على الإطار الزمني المحدد في بقية البرنامج. وكما ناقشنا في بند ٦، فإن واحداً من الأسئلة الأساسية هو ما إذا كان من الممكن تصميم التقييم قبل بدء التدخل، وهذا هو الأفضل. ومن المهم أيضاً معرفة وقت الحاجة للحصول على نتائج التقييم. وفي حال وجود مواعيد نهائية واضحة للحصول على النتائج، على سبيل المثال من أجل تشكيل قرارات حول توسيع نطاق البرنامج أو إصلاحات السياسة، يمكننا التخطيط عكسياً من هذه المعالم لكي نرى إن كنا نملك الوقت الكافي لإجراء أسلوب تقييم الأثر الذي نفكر به. تتطلب بعض الأساليب وقتاً للتنفيذ أطول من أساليب أخرى. يكون للتقييمات المتوقعة (التقييمات المخطط لها مقدماً)، مثل جميع التقييمات العشوائية، بطبيعتها تمتاز بأفاقاً زمنية أطول من التقنيات بالأثر الرجعي، مثل المطابقة البسيطة. يوضح شكل ٧-٢ العوامل الرئيسية التي تؤثر على طول زمن تقييم الأثر. وكما نرى فإن الجدول الزمني للتنفيذ والزمن الضروري من أجل تجسيد الآثار يختلفان من برنامج إلى آخر. وكقاعدة عامة، تستغرق غالباً التقييمات المتوقعة اثني عشر إلى ثمانية عشر شهراً، وتستغرق تقييمات الأثر بأثر رجعي ستة أشهر على الأقل.

شكل ٧-٢ نموذج جدول زمني لتقييم محتمل للأثر

المهمة	ش ١	ش ٢	ش ٣	ش ٤	ش ٥	حتى	ش ١٦*	ش ١٧	ش ١٨	ش ١٩	ش ٢٠
البرنامج											
تصميم البرنامج											
تحديد الفئة المؤهلة											
اختيار المشاركين											
تنفيذ البرنامج											
دمج الدروس المستفادة											
المتابعة والتقييم											
تصميم نظام متابعة**											
تطوير إستراتيجية تقييم الأثر											
تكوين فريق التقييم											
إعداد واختبار أداة المسح**											
إجراء مسح أولي**											
تحليل البيانات الأولية**											
استمرار المتابعة**											
إجراء مسح ختامي**											
تحليل البيانات الختامية											
نشر النتائج											

* يعتمد على الوقت المطلوب لتجسيد الآثار

** ينطبق فقط على التقييمات المتوقعة

عملياً فإن الوقت الأطول الذي تستغرقه التقييمات المتوقعة هو أقل تعقيداً مما يبدو عليه. عندما يتم إعداد برامج جديدة، فإنها تستغرق عدة أشهر لكي تصبح ذات قدرة تشغيلية كاملة. ويمكن القيام بالتحضير لتقييم الأثر خلال مراحل التخطيط للبرنامج ومراحل استطلاع الجدوى ويمكن أن تكون جاهزة بسهولة بحلول الوقت الذي يوشك فيه البرنامج على البدء. وحتى لو كان برنامج ما قد بدء بالفعل وإذا كان قد تم تنظيم البرنامج على مراحل، يمكن التخطيط لإجراء تقييم متوقع للأثر للمرحلة التالية من البرنامج.

الميزانية

يمكن أن تكون تقييمات الأثر مكلفة، ولهذا تتردد العديد من المنظمات في تمويلها. والحقيقة هي أن التكلفة تختلف كثيراً من بلد إلى آخر وحسب المنهجيات والبرامج المحددة التي يتم تقييمها. وتبلغ تكلفة التقييمات عادة من ١٠٠,٠٠٠ دولار وحتى أكثر من ١ مليون دولار. وفي ظروف خاصة للغاية، مثل الحالة التي تتوفر جميع البيانات مسبقاً، يمكن أن تبلغ تكلفة تقييمات الأثر ما يقارب ١٥,٠٠٠ دولار وهو مبلغ صغير. وعند الحاجة لجمع أولي للبيانات، فإنه من غير المحتمل أن تكون تكلفة تصميم تقييم الأثر وتنفيذه أقل من ٥٠,٠٠٠ دولار.

عناصر التكلفة

يرتبط دائماً الأمران الرئيسيان الأكثر تكلفة في تقييم الأثر بالاستشاري ووقت العاملين، وجمع البيانات (أنظر جدول ٧-١).

وقت العاملين. لا يجب الاستقطاع من الوقت المطلوب لاختيار منهجية التقييم الملائمة

وتصميمه. ويمكن عادة لفريق المتابعة والتقييم تصميم التقييم بالتعاون مع استشاري تقييم. ويتراوح أجر الاختصاصي حسب خبرته بين ٢٠٠ دولار إلى ١,٠٠٠ دولار يومياً، وذلك لمدة تصل لعشرين يوماً. وسيكون هناك الحاجة لمزيد من الوقت من أجل تحليل البيانات، والذي يمكن لنفس الاستشاري الذي ساعد في تصميم التقييم القيام به. وعلاوة على ذلك، قد يكون هناك حاجة لمزيد من الاستشاريين من أجل دعم عناصر محددة في التقييم، مثل تصميم الاستبيان. (ستعطي الخطوة التالية، تكوين فريق التقييم، تفاصيل أكثر حول أدوار ومسؤوليات مختلف أعضاء فريق التقييم).

جمع البيانات. إن عنصر التكلفة الرئيسي في أي تقييم للأثر هو بشكل أساسي عملية جمع

البيانات. وتوظيف مؤسسة متخصصة بالمسح هو أمر أكثر تكلفة من جمع البيانات باستخدام طاقم البرنامج، إلا أن هذا التوظيف يضمن جودة أفضل للبيانات. ويعتمد معيار التكلفة لكل شخص تتم معه المقابلة في استبيان أولي على حجم الاستبيان وسهولة إيجاد الأشخاص الذين ستم معهم المقابلة. وفي بعض الحالات، ستبلغ تكلفة استبيان قصير تقوم به مؤسسة متخصصة بالمسح مع أشخاص يمكن تحديدهم بسهولة بمساعدة طاقم البرنامج حوالي ٢٠-٤٠ دولار لكل شخص تتم معه المقابلة. وفي الأماكن التي يصعب فيها النقل أو لا يمكن بسهولة إيجاد الأشخاص الذين ستم معهم المقابلة، يمكن أن تبلغ التكلفة ٥٠-٨٠ دولار لكل شخص تتم معه المقابلة. وتتضمن هذه التكلفة جميع جوانب المسح، بما في ذلك توظيف وتدريب المقابليين، وإجراء المسح، وتقديم البيانات. ويزيد في مسح المتابعة مسائل خاصة تتعلق بتعقب المشاركين وستبلغ التكلفة المحتملة لها ١,٥ مرة من تكلفة المسح الأولية. ومن جهة أخرى وفي حال كان التتبع لا يمثل مشكلة وكانت عينة مجتمع الدراسة مستقرة نسبياً ومن السهل العثور عليها، عندها يمكن أن يكون استبيان المتابعة أقل تكلفة من المسح الأولي. يمكن الاطلاع على طرق لتخفيض التكاليف في ملحق ٢.

تقييم التكلفة

بالنسبة لأغلب التدخلات المتعلقة بسبل عيش الشباب، ربما يكون من الإنصاف أن نفترض بأن إجمالي تكلفة تقييم الأثر ستكون ما بين ١٥٠,٠٠٠-٥٠٠,٠٠٠ دولار. وهذا مبلغ كبير بالنسبة للعديد من البرامج الصغيرة أو متوسطة الحجم، وهو يطرح سؤال ما إذا كانت التكلفة مبررة. تعتمد الإجابة على هذا السؤال بشكل أساسي على (١) الأفق الزمني للبرنامج، و(٢) توقعات النتائج الحالية والمستقبلية. فعلى سبيل المثال، إذا كان الأفق الزمني لبرنامج صغير نسبياً بميزانية سنوية ٢٠٠,٠٠٠ دولار خمس سنوات أو أكثر، أو كان يوجد احتمال لتوسع البرنامج لنقل أنها ستبلغ ٢ مليون دولار سنوياً، عندها يكون إنفاق ٢٥٠,٠٠٠ دولار أمريكي على تقييم لأثر يساهم في

[مصدر إلكتروني]

قائمة بفرص تمويل مختارة

عنوان الموقع الإلكتروني:

<http://www.ifyfnet.org/gpye-m&e-resource4>

إن تقييم تدريب حو الأمية المالية الذي تقدمه فينو FINO في الهند ويتم تنفيذه عبر البنوك المحلية هو مثال عن تقييم يمكن أن تبلغ تكلفته أكثر من البرنامج نفسه. وتبلغ تكلفة تنفيذ البرنامج التدريبي، والذي يستفيد منه حوالي ٣,٠٠٠ مشارك، حوالي ٦٠,٠٠٠ دولار، بينما تبلغ تكلفة التقييم حوالي ٢٠٠,٠٠٠ دولار. وتم تبرير التكلفة على أساس إمكانية التوسع. ولدى البرنامج المصري حالياً أكثر من ٢٥ مليون عميل في الهند وهذا العدد يزيد ٨٠,٠٠٠ شخصاً كل يوم. إن قيمة المعلومات المستقاة من التقييم ليست مفيدة للبرنامج التجريبي فحسب وإنما قد تكون مفيدة أيضاً للملايين المستفيدين المستقبليين.

تصميم البرنامج الموسع هو استثمار مجدي للمال. وفي الواقع، فإن عدم إجراء تقييم للأثر وإدارة مشروع غير فعال يمكن أن يكون أكثر تكلفة بكثير. ومن جهة أخرى، إذا كان من الواضح أن البرنامج نفسه سيستمر لستين فقط، عندها يمكن أن تكون تكلفة التقييم غير متكافئة، على الرغم من أن مجتمع سبل عيش الشباب الأكبر سيستفيد من المعرفة الناتجة عن تلك الدراسة. وفي مثل هذه الحالة، قد يتم اتخاذ القرار اعتماداً على توفر أموال خارجية تنقاسم التكاليف.

شكل ٧-١ نموذج ميزانية تقييم الأثر

مرحلة المتابعة				المرحلة الأساسية				مرحلة التصميم			
التكلفة لكل وحدة	عدد الوحدات	التكلفة الكلية (دولار أمريكي)	الوحدة	التكلفة لكل وحدة	عدد الوحدات	التكلفة الكلية (دولار أمريكي)	الوحدة	التكلفة لكل وحدة	عدد الوحدات	التكلفة الكلية (دولار أمريكي)	الوحدة
٢,٠٠٠	١	٢,٠٠٠	أسابيع	٢,٠٠٠	١	٢,٠٠٠	أسابيع	٤,٠٠٠	٢	٨,٠٠٠	أسابيع
٣,٠٠٠	٣	٩,٠٠٠	أسابيع	٣,٠٠٠	٣	٩,٠٠٠	أسابيع	٣,٠٠٠	٣	٩,٠٠٠	أسابيع
٤,٠٠٠	١٠	٤٠,٠٠٠	أيام	٢,٠٠٠	٥	١٠,٠٠٠	أيام	٤,٠٠٠	١٠	٤٠,٠٠٠	أيام
١,٥٠٠	٥	٧,٥٠٠	أيام	٠	٠	٠	أيام	١,٥٠٠	٥	٧,٥٠٠	أيام
١٠,٠٠٠	١٠٠	١,٠٠٠,٠٠٠	أيام	٨,٠٠٠	٨٠	٦٤٠,٠٠٠	أيام				
٦,٠٠٠	٢	١٢,٠٠٠	رحلات	٦,٠٠٠	٢	١٢,٠٠٠	رحلات	٦,٠٠٠	٢	١٢,٠٠٠	رحلات
٧٥٠	٥	٣,٧٥٠	أيام	٧٥٠	٥	٣,٧٥٠	أيام	٧٥٠	٥	٣,٧٥٠	أيام
٦,٠٠٠	٢	١٢,٠٠٠	رحلات	٦,٠٠٠	٢	١٢,٠٠٠	رحلات	٦,٠٠٠	٢	١٢,٠٠٠	رحلات
٣,٠٠٠	٢٠	٦٠,٠٠٠	أيام	٣,٠٠٠	٢٠	٦٠,٠٠٠	أيام	٣,٠٠٠	٢٠	٦٠,٠٠٠	أيام
١٢٠,٠٠٠	٢,٠٠٠	٢٤٠,٠٠٠	الشباب	٨٠,٠٠٠	٢,٠٠٠	١٦٠,٠٠٠	الشباب				
إجراء مسح											
هـ. النشر											
٥,٠٠٠	١	٥,٠٠٠	النشر/ الطباعة								
٥,٠٠٠	١	٥,٠٠٠	الورشات								
١٦٦,٢٥٠				١١٠,٧٥٠				٢٨,٢٥٠			
إجمالي التكلفة لكل مرحلة											
٣٠٥,٢٥٠											
إجمالي تكلفة التقييم											

* تتضمن التدريب، والاختبار، ومواد المسح، والطاقم الميداني (المقايين والمشرفين)، والنقل، وغيرها. المصدر: مقتبس من غيرتلر وآخرين (٢٠١١).

تكوين فريق التقييم

تتطلب تقييمات الأثر مجموعة من المهارات التي بدورها تستلزم في العادة فريق تقييم واسع يضم من جهة المسؤولين عن البرنامج الذين سيحددون ما إذا كان من حاجة إلى إجراء تقييم للأثر ويصيغون أسئلة التقييم ويشرفون على جهود التقييم بشكل عام ومن جهة أخرى، خبراء التقييم الذين عادة ما يكونون مستشارين ويكونون مسؤولين عن أوجه التقييم التقنية، بما فيه اختيار المنهجية الملائمة والتخطيط لجمع البيانات وإجراء التحليل.

يتألف الفريق الأساسي من مدير البرنامج وموظف المتابعة والتقييم (كلاهما من الداخل)، وخبير تقييم رئيسي (وغالباً ما يدعى الباحث الرئيسي أو PI)، ومساعد في البحث يعمل مع الباحث الرئيسي ولتصاميم التقييم التي تتضمن جمع بيانات جديدة، يضاف إلى الفريق خبير في المسح ومنسق ميداني وفريق العمل الميداني (مثل مؤسسة متخصصة في جمع البيانات) بالإضافة إلى مدراء البيانات ومعالجتها. يعرض جدول ٧-٢ أدوار كل شخص ومسؤولياته. استناداً إلى حجم البرنامج والتقييم ومستوى مهارات أفراد الفريق، يمكن توكيل عدة مهام لشخص واحد.

[مصدر إلكتروني]

المعطيات اللازمة لطاقم عمل تقييم الأثر الرئيسي

عنوان الموقع الإلكتروني:
<http://www.iyfnet.org/gpye-m&e-resource10>

من	المهام الكبيرة	الشخصية/المهارات المطلوبة
مدير المشروع	• تحديد أهداف التعلم • تقدير الموارد اللازمة • تحضير الشروط المرجعية للباحث الرئيسي • توظيف استشاري التقييم	• خبرة في تصميم برامج سبل عيش الشباب وتطبيقها • خبرة في إدارة فريق • قدرة على إعداد الميزانيات • قدرة على العمل عن كثب مع فرق البرنامج والتقييم
موظف/ وحدة المتابعة والتقييم الداخلية	• تحديد النموذج النظري للبرنامج (سلسلة النتائج) • تحديد المؤشرات وأدوات القياس • إدارة جهاز المتابعة عند بدء البرنامج	• شهادة ماجستير في الاقتصاد أو السياسة العامة أو مجال ذي علاقة • قدرة على العمل عن كثب مع فرق البرنامج والتقييم • قدرة على الموازنة بين المتابعة ومسؤوليات تقييم الأثر
الباحث الرئيسي (جامعة محلية أو دولية، مركز فكري بحثي، استشارة متخصصة)	• اختيار تصميم التقييم • تعديل التصاميم الجيدة نظرياً لتناسب الميزانية والوقت والبيانات والقيود السياسية الواقعية • تطوير مقاربات بمزج طرق عدة • تحديد فريق التقييم وتحضير المعطيات • الإشراف على طاقم العمل • تحديد متطلبات المعاينة والقوة • تحليل البيانات وصياغة التقرير	• شهادة ماجستير في الاقتصاد أو السياسة العامة أو مجال ذي علاقة • معرفة للبرنامج أو لأنواع مشابهة من البرامج • خبرة في أساليب البحث والتحليل الاقتصادي القياسي • بعض الخبرة في البلد أو المنطقة المعنية • قدرة مثبتة على العمل بفعالية في فرق متعددة التخصصات • مهارات شفوية وكتابية عالية
خبير في الاستبيان (قد يكون الباحث الرئيسي نفسه)	• تصميم أداة المسح • تحضير الكتيبات المرفقة وكتب الترميز • تدريب مؤسسة جمع البيانات • دعم اختبار الاستثمارات ومراجعتها	• شهادة البكالوريوس أو الماجستير في الاقتصاد أو السياسة العامة أو حقل مرتبط • خبرة في إجراء المسح على الأطفال والشباب • خبرة في إجراء العمل الميداني في البلد أو المنطقة المعنية • قدرة على التفاعل بشكل فعال مع النظراء في البحث والبرنامج
المنسق الميداني وفريق العمل الميداني	• المساعدة في إعداد الاستمارة • توظيف وتدريب المقابليين • تشكيل فرق العمل الميداني وتحديد الجدول الزمني لعملها • الإشراف على جمع البيانات • تنقية البيانات للتمكن من مشاركتها مع اختصاصي التقييم	• أهلية قانونية، رخص للعمل تعترف بها حكومة البلد حيث يتم العمل • شبكة جيدة من المقابليين والمشرفين ومدخلي البيانات المحترفين • خبرة مثبتة من خمس سنوات أو أكثر في تنظيم المسوح على نفس حجم البرنامج • قدرة قوية على التخطيط وتنظيم التجهيزات للمسح • قدرة قوية على إدارة البيانات والإحصائيات • القدرة على السفر والعمل في ظروف صعبة
مساعد في البحث	• تحليل البيانات • دعم الباحث الرئيسي في صياغة تقارير التقييم	• شهادة البكالوريوس أو الماجستير في الاقتصاد أو السياسة العامة أو مجال ذي علاقة
مدراء ومعالجو البيانات	• تنقية البيانات ليتمكن مساعد البحث والباحث الرئيسي من استخدامها • إدارة فريق البيانات	• خبرة في برمجيات البيانات وإدارة فريق البيانات

[مصدر إلكتروني]

مصادر لإيجاد خبراء في تقييم الأثر

عنوان الموقع الإلكتروني:

<http://www.iyfnet.org/gpye-m&e-resource5>

[معلومة إرشادية]

غالباً ما يعتبر عقد الشراكات مع المؤسسات الأكاديمية استراتيجية قوية للمنظمات غير الحكومية والحكومات لتطوير قدراتها على إجراء تقييمات الأثر. على سبيل المثال:

- عقدت مؤسسة إنقاذ الأطفال شراكة مع جامعة لوس آنديس في كولومبيا لتقييم مبادرة يوث سايف (YouthSave).
- عقدت منظمة شباب الأعمال الدولية ولجنة تطوير الأرياف في بنغلاديش BRAC شراكة مع كلية لندن للاقتصاد.
- عقدت وزارة العمل التركية شراكة مع جامعة الشرق الأوسط التقنية لتقييم وكالة التوظيف العامة التركية (ISKUR).

بعد تصميم التقييم الأساسي وجمع البيانات الأولية، وحالما يبدأ البرنامج، لن يكون أمام مدير المشروع وموظف المتابعة والتقييم الكثير ليفعله. لذا، من الصائب جعل واحد منهما، ربما موظف المتابعة والتقييم، يعمل على التقييم بدوام جزئي في خلال هذه الفترة للحرص على المتابعة الملائمة للبرنامج. وإذا طرأت أية مشاكل كبيرة على تطبيق البرنامج، سيتوجب توثيق ذلك وفي بعض الحالات الإبلاغ عنها للفريق الأشمل.

ولا يجب توظيف كل الخبراء من الخارج في ذات الوقت. إذ الأمر الأهم هو اختيار الباحث الرئيسي الذي عليه أن يبقى طوال مدة التقييم، أي منذ تصميم التقييم وحتى صياغة التقرير النهائي، للحرص على الاستمرارية (مع أنه من المرجح أنه لن يعمل على التقييم خلال فترة تطبيق البرنامج). ويمكن اختيار أعضاء آخرين من الفريق إضافة إلى المقيم الرئيسي عند الحاجة. يتم اللجوء عادة إلى خدمات خبير تطوير المسح على سبيل المثال لمهام قصيرة وقد لا يشارك في التقييم إلا لأسابيع قليلة، وذلك استناداً إلى حجم التقييم. كما يتم توظيف مؤسسة جمع البيانات لإجراء المسح الأولي والنهائي ويفضل أن تكون المؤسسة نفسها معنية بكلا مرحلتي جمع البيانات، مع أن ذلك لا ضرورة أو جدوى له في كل الحالات..

وضع خطة تقييم

حالما يتم تعيين الباحث الرئيسي، سيحضر هذا الأخير عادةً خطة لتقييم الأثر (تعرف أيضاً بورقة المفاهيم) بالتنسيق مع قادة البرنامج. وعلى هذه الخطة أن تصف أهداف التقييم وتصميمه واختيار عيناته واستراتيجيات جمع بياناته. ستكون خطة تقييم الأثر، بشكل عام (أنظر إلى موجز نموذجي في مربع ٧-١)، أساس منهجية تقييم الأثر التي سيتم اختيارها وسترشد جميع الخطوات التالية في عملية تطبيق التقييم.

مربع ٧-١ موجز لخطة تقييم أثرالمطابقة

١. مقدمة
٢. خلفية
٣. التدخل
٤. تصميم التقييم
- ٤-١ هدف التقييم
- ٤-٢ الفرضيات وأسئلة البحث
- ٤-٣ أسلوب التقييم
٥. إستراتيجية وقوة اختيار العينات
٦. خطة جمع البيانات
٧. خطة تحليل البيانات
- ٧-١ قياس الأثر
- ٧-٢ فحص تأثيرات التفاضل
- ٧-٣ قياس عائد البرنامج (تحليل المنافع والتكاليف)
٨. المخاطر والتدابير المقترحة للتخفيف من وطأتها
٩. الجمهور والنشر
١٠. الجدول الزمني والنشاطات
١١. الميزانية
١٢. المرفقات

كل بين مثال تخطيط تقييم أثر القروض الصغيرة للشباب في اليمن أهمية التعاون الوثيق فيما بين طاقم البرنامج والمقيمين منذ بدء البرنامج للحرص على فهم الطرفين للسياق التشغيلي. صمم المقيمون على حدة تجربة ضبط عشوائية لتقييم أثر القروض والخدمات المالية الأخرى للشباب على خلق فرص العمل وتوسيع الأعمال ونتائج أخرى. وعندما حان وقت عرض تصميم التقييم، وضح المدير العام للمصرف أن مثل ذلك التصميم كان غير مقبول في سياق مؤسسة مالية حديثة التأسيس لا يمكنها أن تتحمل تكلفة استبعاد زبائن محتملين لأجل غرض التقييم. وتوجب على فريق التقييم عندئذ أن يبدأ من الصفر واختار في النهاية تصميمًا لتقييم ترويج عشوائي تبين أنه أكثر ملائمة لتدخل ذي تغطية عالمية.

لا يجب أن يتم تطوير تصميم التقييم (النقطة الرابعة) على يد خبير التقييم على انفراد؛ بل يجب أن تشرك العملية طاقم البرنامج بشكل وثيق، للحرص على أن طريقة التقييم المختارة تناسب أهداف التعلم وسياق البرنامج التشغيلي (أنظر بند ٦ لمناقشة مفصلة للموضوع). إضافة إلى ذلك، بالرغم من أن الباحث الرئيسي سيتوجه دون شك إلى طاقم البرنامج ويقدم اقتراحات لتحديد عينة التقييم (النقطة ٥) والتخطيط لجمع البيانات (النقطة ٦)، يبقى من المفيد أن تكون المنظمة المطبقة على اطلاع حول صلة هذه الأوجه بالتقييم وبالبرنامج نفسه. لذلك، نتطرق إلى هاتين النقطتين بمزيد من التفاصيل أدناه.

تحديد عينة التقييم

لا نحتاج بالضرورة إلى تقييم كل مشترك في البرنامج لتقييم التدخل. ليس علينا سوى اختيار مجموعة ممثلة للأشخاص - عينة - تكون كبيرة بما يكفي لغرض تقييمنا. إذا كانت عينتنا تمثل كل الشباب المؤهلين، يمكننا عندئذٍ تعميم نتائج تقييمنا على مجموع السكان المؤهلين. فنحن بحاجة لأن تتمتع النتائج بصدق خارجي بالإضافة إلى الصدق الداخلي النابع من بناء مجموعة مقارنة جيدة. وللحصول على عينة ممثلة، نحتاج إلى استراتيجية ما في اختيار العينات.

يجب أن تكون العينة أيضاً كبيرة بما يكفي لتتمكن من توليد مقارنة موثوقة للنتائج بين المشاركين في مجموعة المعالجة وأولئك في مجموعة المقارنة. إذا كانت العينة صغيرة جداً، قد لا تتمكن من رؤية أثر ذا دلالة إحصائية للبرنامج، وحتى إن وجد هذا الأثر بالفعل. ولمعرفة كم يجب أن يكون حجم العينة ليكون كبيراً بما يكفي، نحتاج إلى القيام بحسابات القوة. تمت مناقشة هذه المفاهيم أدناه.

وضع استراتيجية اختيار العينة

تشمل استراتيجية اختيار العينة الخطوات الثلاث التالية:

١. تحديد مجتمع الدراسة موضع الاهتمام. أولاً: نحتاج إلى تكوين فكرة شديدة الوضوح عن المستهدفين والمؤهلين للالتحاق بالبرنامج. على سبيل المثال، يمكن لبعض العوامل مثل العمر والجنس ومستوى الدخل والوضع الوظيفي والموقع أن تحدد مدى الأهلية. لن تشمل الدراسة هؤلاء غير المؤهلين.
٢. تحديد إطار العينات. يعتبر إطار اختيار العينات هو القائمة المتاحة الأكثر شمولاً للفئة المستهدفة موضع الاهتمام، حيث يتيح لنا معرفة الصلة بين العينة التي حصلنا عليها وبين الفئة موضع الاهتمام بوجه عام ممن نرغب في استخلاص دروس التقييم لأجلهم. من وجهة النظر المثالية، يتطابق إطار العينة تماماً مع الفئة موضع الاهتمام، مما يجعلهم ممثلين مثاليين لها، بينما نجد من وجهة النظر العملية أننا نحاول الحصول على قائمة من الشباب المؤهلين من سجلات مسح السكان أو المدرسة أو سجلات الناخبين أو سجلات المدينة التي تحوي بلا شك عدد هائل من الشباب المؤهلين. من الصعوبة عملياً الحصول على إطار لاختيار العينات التي تغطي بشكل كامل كل الفئة المؤهلة.
٣. سحب عدد الوحدات المنشود من إطار العينات باستخدام أحد أساليب اختيار العينات المتاحة. يمكن استخدام العديد من الأساليب لسحب العينات من الإطار المحدد لدينا، أما الأساليب الأكثر شيوعاً فهي بعض أشكال أخذ عينات احتمالية. من خلال هذا الأسلوب، يتم اختيار المشاركين داخل العينة باحتمالية محددة. في حالة الاختيار العشوائي للعينة، على سبيل المثال، كان لكل مشارك في إطار العينة احتمالية الالتحاق. وعند استخدام إجراءات أخذ عينة غير احتمالية، فإننا نواجه خطورة الحصول على عينة لا تمثل الفئة المؤهلة عموماً.

[تعريف]

الصدق الخارجي: قدرتنا على تعميم النتائج. يعود ذلك إلى أي مدى يمكننا توقع الحصول على النتائج نفسها إذا قدمنا البرنامج إلى مجموعة مختلفة أو ذات حجم أكبر. لضمان ذلك، نحتاج إلى استراتيجية ملائمة لاختيار عينة الأفراد الذين سنعمل معهم.

إطار العينات: قائمة الوحدات الأكثر شمولية التي يمكننا الحصول عليها للفئة المعنية. يمكننا سحب الأسماء من هذه القائمة بأن نحصل على عينة.

العينة: العينة هي مجموعة ضمن الفئة المستهدفة. عادة ما يكون من المستحيل ومن غير العملي جمع المعلومات عن الفئة المستهدفة كلها. يمكننا عوضاً عن ذلك جمع المعلومات عن مجموعة ضمن الفئة تكون ذات حجم يمكن إدارته. وإذا تم اختيار التركيبة بشكل جيد، من الممكن عندئذٍ القيام باستنتاجات أو استخلاصات عن جميع الأفراد في تلك الفئة.

حسابات القوة، أو «ما هو الحجم الذي يجب أن تكون عليه العينة التي اختارها؟» إنه لمن المهم معرفة الحجم المثالي للعينة التي نختارها، أو بمعنى آخر، معرفة ما هو عدد الأفراد الذين ينبغي علينا سحبهم من إطار العينة. إذا كانت العينة صغيرة جداً، قد يقودنا التحليل الإحصائي لها إلى الاستنتاج بأن البرنامج لم يحقق أية آثار إيجابية للمستفيدين، بينما يكون ذلك مخالفاً للواقع. وفي المقابل، فإن جمع كم هائل من المعلومات على نحو يفوق اللازم يعد أمراً مكلفاً للغاية. تساعدنا حسابات القوة على تقدير الحجم الصحيح للعينة حيث تدل على أصغر عينة والتي يمكننا من خلالها فعلياً قياس أثر برنامجنا بمستوى معقول من الثقة.

على الرغم من تفاوت الحجم المناسب لعينة التقييم، يتعين علينا بوجه عام أن نشمل ما بين ١,٠٠٠ - ٣,٠٠٠ شاب في تقييمنا من أجل ضمان وجود عدد وافٍ من الشباب لكل من مجموعة المعالجة ومجموعة المقارنة. في بعض الحالات المحدودة جداً، يمكن الاكتفاء بعينة تضم أقل من ١,٠٠٠ مشارك، غير أنه لا ينصح بالعينات التي تقل عن ٥٠٠ مشارك (٢٥٠ مشارك لمجموعة المعالجة و٢٥٠ آخرين لمجموعة المقارنة). يستطيع خبراء التقييم الاضطلاع بحساب الحجم المناسب للعينة للتقديم الذي من المقرر إجراؤه.

تخطيط جمع البيانات

تحتاج خطة التقييم إلى وضع استراتيجية لجمع البيانات. عادةً ما تكون عملية جمع البيانات مهمة معقدة للغاية ولهذا يفضل إسنادها إلى فريق من الخبراء خارج الفريق. تشمل القضايا الرئيسية هنا على بعض الموضوعات كتوقيت جمع البيانات وهل يجب جمع بيانات جديدة ومن الذي سيتولى جمعها وكذلك كيف سيتم إدارة البيانات. نقوم فيما يلي بمناقشة هذه القضايا.

توقيت جمع البيانات

يعد توقيت جمع البيانات في غاية الأهمية ويعتمد على طبيعة البرنامج. عند استخدام المسح الأولي، يجب إتمامه قبل بداية البرنامج وكذلك قبل معرفة المشاركين بقبول التحاقهم بالبرنامج لضمان ثبات اجاباتهم في مجموعتي المعالجة والمقارنة، وهو أمر ضروري للغاية، إذ قد يغير الشباب في إجاباتهم إذا علموا بأمر قبولهم في البرنامج.

كذلك ينبغي عند تحديد توقيت مسح المتابعة الأخذ في الاعتبار حاجات البرنامج وآثاره. فإذا قمنا بعمل مسح المتابعة في وقت قريب جداً، لن نلاحظ أية آثار مرتبطة على البرنامج، بينما إذا قمنا بعمل هذا المسح في وقت لاحق عما يجب القيام به، لن يستفيد البرنامج مما توصل إليه المسح من معلومات.

البيانات الكائنة مقابل البيانات الجديدة

ليس من الضروري دائماً جمع بيانات جديدة. في بعض الحالات، تكون البيانات المطلوبة لإجراء تقييم موحدة (يعرض مربع ٧-٢ اقتراحات عن كيفية العثور على هذه البيانات). يوجد هناك نوعان من البيانات الشائعة ويجب استكشافها قبل اتخاذ القرار بجمع بيانات جديدة.

[تعريف]

القوة هي احتمالية اكتشاف أثر ما عند ظهوره. لكن تبقى هناك دائماً مخاطر أننا لن نكتشف بعض الآثار رغم وجودها. على الرغم من ذلك، في حالة الانخفاض الشديد لاحتمالية عدم اكتشاف أثر ما موجود، يمكننا وقتئذ القول بأن الدراسة قوية بشكل كافٍ.

[مصدر إلكتروني]

مثال على تقدير حجم العينة

عنوان الموقع الإلكتروني:
<http://www.iyfn.net/org/gpye-m&e-resource6>

[معلومة إرشادية]

نتيجة احتمالية انسحاب بعض المشاركين من البرنامج أثناء التنفيذ وخروجهم بذلك من التقييم، يوصى باختيار عينة أكثر عدداً من الحد الأدنى للعينة التي تشير إليها حسابات القوة.

في أحد البرامج المنفذة بالاشتراك مع الحكومة المحلية، واجهت إحدى المنظمات غير الحكومية بأمريكا اللاتينية تأخيرات متعددة أثناء اختيار المشاركين، حيث انقضى الكثير من الوقت بين اختيار الشباب المشاركين وبين بداية التدريب، وبدأ المشاركون يفقدون اهتمامهم بالمشاركة وانسحبوا من مجموعة المعالجة، مما أدى إلى انخفاض عدد المشاركين. مجموعة المعالجة عن العدد المناسب. في مثل هذه الحالات، ينبغي أن يكون الأثر كبير جداً حتى يكون قابلاً لقياسه.

البيانات الإدارية. يتم في الغالب جمع البيانات الإدارية من خلال برنامج تنفيذي لأغراض المراقبة.

بيانات مسح الأسر. يتم إجراء مسح للأسر على المستوى القومي بشكل دوري في العديد من الدول النامية. وتتضمن مسح حول موضوعات متعددة مثل مسح قياس مستويات المعيشة ومسح الديموغرافيا والصحة والتي يمكن أن تشمل نطاقاً واسعاً من المعلومات حول خصائص الإسكان واستهلاك الأسر وثرواتهم والمؤشرات الخاصة بتوظيف الأفراد والتعليم والصحة. تكون بعض المسوح الأخرى مثل مسح قوى العمل أكثر تحديداً في النطاق وبعض الأحيان تغطي المناطق الحضرية فقط.

أماكن البحث:

- المعاهد الإحصائية في البلد ذات الصلة
- الشبكة الدولية لمسوح الأسر (www.ihnsn.org)
- مسح الديموغرافيا والصحة (<http://www.measuredhs.com/>)
- مسح قياس مستويات المعيشة (<http://iresearch.worldbank.org/lsms/lsmssurveyFinder.htm>)

بيانات التعداد السكاني. تجري معظم الدول تعداداً للسكان والأسر كل عشر أعوام، كما تجري العديد من هذه الدول المزيد من المسوح الإضافية. ومن أهم مميزات بيانات التعداد السكاني هي أنها تشمل السكان برمتهم، لذلك فهي تحوي بيانات تتيح مقارنات ومعالجات فعلية. والنقطة السلبية في تعدادات السكان هو أنها قليلة التكرار وتحتوي على عدد محدود من المؤشرات الأمر الذي يحد من قيمتهم لتقييم الأثر.

أماكن البحث: الشبكة الدولية لمسوح الأسر (www.ihnsn.org)

بيانات مسح المنشآت: تقوم مسح المنشآت بجمع البيانات على مستوى تقديم الخدمة، على سبيل المثال في المدارس أو مراكز التدريب المهني. وقد تعمل الوزارات المحلية أو أجهزة الدولة أو حتى الهيئات المحلية على جمع المعلومات. وفي العديد من الحالات، توفر المسوح التي تجري على مستوى المنشآت متغيرات ضابطة (مثل نسبة المدرس إلى الطالب)، بينما قد تقدم مسح أخرى نتائج ذات أهمية تخصها مثل معدلات الحضور والغياب.

أماكن البحث: وزارت محلية ذات الصلة والممثلين المحليين.

بيانات مسح متخصصة. المسح المتخصص هو ذلك الذي يتم جمعه لغرض محدد، وفي الغالب يكون لإجراء بحث حول موضوع بعينه. يستعين العديد بنماذج من المسح القومي للأسر ويضيف أسئلة حول الموضوعات التي تهتم. قد يكون نطاق التغطية في المسوح المتخصصة محدوداً، وفي بعض الأحيان ينتج عنه تداخل ضئيل أو عدم تداخل مع إهتمامات البرنامج. إلا أنه، إذا تمكن فريق التقييم من العثور على البيانات الكائنة من المسح المتخصص حول موضوع الذي سيجري له التقييم، يمكن لمجموعة البيانات هذه توفير مجموعة غنية من المؤشرات ذات الصلة.

أماكن البحث: المسؤولون المحليون والمانحون والمنظمات غير الحكومية في المنطقة المعنية.

المصدر: مقتبس من البنك الدولي (٢٠٠٧، ص ٨-١١)

[مصدر إلكتروني]

بروتوكول استخدام إحدى الشركات المختصة بالمسوح عنوان الموقع الإلكتروني <http://www.iyfnet.org/gpye-m&e-resource7>

[معلومة إرشادية]

في بعض الحالات، تحاول البرامج أن يتم جمع البيانات من قبل الشركاء المنفذين من خلال العاملين في البرنامج لديهم. لا ينصح باللجوء إلى أشخاص معتمدين على التمويل لإجراء عملية جمع البيانات ويعزى ذلك إلى وجود فرصة أكبر في انحياز النتائج لصالح البرنامج. إذا تم اتخاذ القرار بإجراء عملية جمع البيانات داخلياً، فالأفضل تولى فريق منفصل القيام بهذه العملية يكون من شأنه التركيز فقط على جمع البيانات ولا يكون مرتبطاً بالبرنامج.

[مصدر إلكتروني]

أدوات جمع البيانات القائمة على تكنولوجيا المعلومات والاتصالات عنوان الموقع الإلكتروني <http://www.iyfnet.org/gpye-m&e-resource2>

أولاً، قد تكون البيانات الضرورية تم جمعها بالفعل في شكل بيانات إدارية للمتابعة والتقييم. بناءً على الأسئلة التي يرغب البرنامج في الحصول على إجابة لها، ربما تكون الإجابات قد تم جمعها بالفعل. على سبيل المثال، في العديد من برامج سبل العيش، يتم طلب معلومات حول الدخل والوظيفة عند بداية البرنامج، وبالتالي تحد من الحاجة إلى تقييم أولي. إلا أن هذه المعلومات يتم جمعها فقط

لهؤلاء الملحقين بالبرنامج. كما يجب جمع البيانات حول الأفراد في مجموعة المقارنة. ومن أجل تجنب الانحياز غير المقصود من خلال جمع للبيانات بأسلوب غير متناسق، من المهم أن يكون أي نظام يتم تصميمه لجمع البيانات متناسقاً وموضوعياً بقدر الإمكان لكل من مجموعتي المعالجة والمقارنة. الأمر الذي يصعب تطبيقه من خلال عملية جمع بيانات إدارية بحتة. وإنه لمن الأفضل استخدام فريق مخصص لجمع البيانات الجديدة عن كل من مجموعتي المعالجة والمقارنة، إلا إذا كان هذا نظام جمع المعلومات بطبيعة الحال جزءاً من البرنامج.

ثانياً، دائرة/مكاتب الإحصاءات المحلية ربما يكون لديها بيانات تم جمعها بالفعل حول العديد من المشاركين في البرنامج ومجموعات المقارنة. بالنسبة إلى البرامج الصغيرة، يندر أن نجد عدداً كافياً من الأشخاص الملحقين بالبرنامج كانوا جزءاً من مسح قائم. أما بالنسبة للبرامج الكبيرة، من المحتمل أن نجد بعض المشاركين بها كانوا جزءاً من مسح ما. فضلاً عن ذلك، من الضروري أيضاً فهم البيانات التي تم جمعها وكيفية جمعها. كما أن التأكد من أن الأسئلة المطروحة تتعلق بالبرنامج الذي نريده وأن حجم العينة كبير وكفي لضمان استخلاص الاستنتاجات. تحقق من دائرة/مكاتب الإحصاءات المحلية للتأكد من وجود البيانات وإمكانية استخدامها. إذا لم يكن استخدام المعلومات الكائنة كاف، يجب جمع بيانات جديدة.

فريق جمع البيانات الخارجية مقابل الداخلية

تعد عملية جمع البيانات الجزء الأكثر تكلفة في التقييم ويعزى ذلك إلى أسباب واضحة. إن جمع بيانات عالية الجودة والتي يمكن تحليلها بسهولة هي أساس الوصول إلى تقييم ناجح. فدون الحصول على بيانات عالية الجودة، قد يضيع العمل على تصميم التقييم هباءً. فعند اتخاذ القرار بين استخدام مؤسسة للدراسات المسحية أو تولي فريق العمل الداخلي عملية جمع البيانات، ينبغي على البرنامج اختيار النهج الذي يتناسب مع ميزانيته ويضمن في الوقت عينه جمع بيانات منهجية ذات جودة. تتجه بعض البرامج إلى جمع البيانات بنفسها لتوفير المال. يمكن إتباع هذا النهج في المسوح القصيرة والبسيطة، إلا أنه يكون لها بعض السلبيات الهامة، لا سيما في عمليات جمع البيانات المكثفة. ولا ينصح عادة في تولي العاملين بالبرنامج عملية جمع البيانات لما تنسم به هذه العملية من تعقيد. وفي حين أن استخدام مؤسسة للدراسات المسحية أكثر تكلفة من جمع البيانات من قبل العاملين داخلياً، وهذا أيضاً يعني أنه يمكن جمع البيانات على نحو أسرع وبجهد أقل من قبل المكتب المختص بعمل البرنامج. كما يضمن كذلك جمع البيانات من قبل فريق مؤهل. (المزيد من الإرشادات لضمان الجودة تحت قسمي تدريب فريق العمل الميداني والإشراف على جمع البيانات، أدناه). بالإضافة إلى ذلك، إن استخدام مؤسسة خارجية يضمن الحيادية ويزيد من مصداقية نتائج التقييم.

عملية جمع البيانات وأسلوب جمعها

بوجه عام، يجب أن يدير المسوح موظفين مدربين؛ ويجب ألا يتم استخدام الاستبيانات الذاتية إلا في ظروف معينة. عندما يملأ الأفراد المسوح بأنفسهم كثيراً ما يفسرون الأسئلة بشكل مختلف عما كان يقصده فريق المسح. يعمل الأشخاص الذين يجرون المقابلات المدربين على ضمان وجود اتساق أكبر في التفسير، وكذلك في العديد من السياقات، لا يكون المشاركون متعلمين أو مثقفين كما قد نتوقع أو نأمل لذا فقد يحتاجون إلى مقابلات توجيهية.

توجد عدة طرق لجمع وتسجيل استجابات المسح. تعد المسوح الورقية تقليدية، وإذا أمكن يمكن أن يستخدم الأشخاص الذين يجرون المقابلات الهواتف الخلوية كذلك (حيث يمكن تحميل برامج المسوح عليها) أو أجهزة الكمبيوتر أو المساعدات الرقمية الشخصية. قد يكون من الممكن كذلك التسجيل الصوتي لاستجابات الأشخاص الذين سيتم إجراء المقابلات معهم. على الرغم من أن الأدوات القائمة على التكنولوجيا قد تتطلب بعض التدريب المبني (عادة يكون صغير نسبياً) إلا أنها قد تقلل الوقت اللازم لكل مقابلة، وتخفف الوقت اللازم لإدخال البيانات، وتقلل أخطاء البيانات التي تنشأ عن عملية الإدخال التقليدية للبيانات ومعالجتها، ويمكنها من ثم توفير الوقت والمال خاصة في المسوح الأكبر. من ناحية ثانية، يحتاج الفرد إلى دراسة مدى ملاءمة استخدام المعدات المكلفة أحياناً مع الأسر والأحياء الفقيرة.

وضع واختبار أداة المسح

إذا استدعت خطة التقييم جمع بيانات جديدة من المهم اختيار أداة جمع البيانات الصحيحة. في معظم الحالات، سيتم استخدام نوع ما من أنواع المسوح وكثيراً ما يكون مقترناً بأساليب نوعية أخرى مثل مجموعات النقاش البؤرية/المركزة أو مقابلات المبلغين الأساسيين. بما إن المسح سيكون أساساً لجمع البيانات عن المشاركين وعن مجموعة المقارنة، يعد تصميم المسح أمراً هاماً. على الرغم من أن تصميم الاستبيانات قد يبدو أمراً تافهاً إلا أن الإتيان بمسح عالي الجودة ينتج عنه نتائج موثوقة بها يعد علماً وفناً. إن إجراء المسوح على المراهقين والشباب يثير تحديات إضافية مقارنة بإجراء مسوح على البالغين لذا قد يكون من الحكمة السعي للحصول على دعم من استشاريين خبراء لاتخاذ هذه الخطوة (أنظر مربع ٧-٣).

مربع ٧-٣ العوامل المؤثرة على موثوقية البيانات أثناء إجراء المسح على الشباب

يعتمد أي تقييم على المعلومات الموثوقة. وبينما يشير البحث إلى أن الشباب بوجه عام يكونون مستجيبين موثوق بهم إلا أن هناك عدد من الأسباب التي قد تجعل الشباب يدلون بإجابات خاطئة أو حتى يحرفون إجابات الأسئلة:

- **الفهم.** قد يكون الشباب أقل تعلماً ولديهم قدرة إدراكية محدودة. هل يفهم المستجيب السؤال؟ هل تمت صياغة السؤال بلغة مناسبة للعمر؟ بعض الأسئلة تكون متقنة وقد يصعب على الشباب فهمها حتى عند توجيهها بشكل بسيط ومباشر.
- **التذكر.** إلى أي مدى يمكن أن يتذكر المستجيب الأحداث أو المعلومات؟ إن لهذا علاقة جزئية بفترة المرجعية: منذ متى حدث الحدث أو كم المرات التي تكرر فيها. بوجه عام، تكون فترات التذكر الأقصر أدق من الفترات الأطول.
- **السرية.** هل يوجد أي سبب لدى المستجيب ليشي الانتقام أو أية تداعيات أخرى ناشئة عن الإجابات التي يدلي بها؟ هل يتم إجراء المقابلة بالفعل على انفراد وفي جلسات خاصة؟ يجب أن يكون الشخص الذي يجري المقابلة قادراً على إقناع المستجيب بأن المعلومات ستعامل بسرية.
- **القبول الاجتماعي.** هل يعتقد المستجيب بأن الشخص الذي يجري المقابلة يتوقع استجابة ما دون غيرها؟ هل يمكن رؤية إجابة ما على أنها «صحيحة»؟ يؤثر هذا بشكل خاص على السلوكيات غير القانونية أو الموصومة بالعار أو الخاضعة لقيود أخلاقية. تظهر دراسات تقرير بيرنرويللي وجريدي (٢٠٠٣) أنه يرجح أن يبلغ المراهقون عن تعاطي الكحوليات حديثاً في الاستبيانات الذاتية أكثر من المقابلات في حين أنه لا يوجد اختلاف في استجابات البالغين. بالإضافة إلى ذلك، أكدت العديد من الدراسات أن الشباب على الأرجح يقدمون إجابات متناقضة أكثر من البالغين في المسوح المكررة على مدار الوقت.
- **الإرهاق.** بالرغم من أن المسوح بين البالغين يمكن أن تستغرق ساعات لاستكمالها إلا أن الشباب هم الأكثر قابلية لأن ينفذ صبرهم في المقابلات الطويلة. على سبيل المثال، مؤسسة إنقاذ الأطفال وهي منظمة غير حكومية أنشأت مؤشر تنمية سبل عيش الشباب ويتألف من ثلاثة مسوح مدارية ذاتياً للشباب في أعمار من ١١-٢٤ عاماً لاستخراج معلومات عن الأصول والكفاءات. وجد الاختبار التجريبي أن الشباب «قد أصابهم الملل من الاستبيانات الطويلة وو قدموا إجابات مغبركة» (بيرتراند وآخرون، ص ٥).

مجموعة أدوات المسح متضمنة الاستبيانين الأساسي والختامي لصندوق العمل الاجتماعي لشمال أوغندا.
عنوان الموقع الإلكتروني

ملاحظة: إن استبيان NUSAF طويل للغاية وبرغم أنه كان قائماً على مسح سابق إلا أنه استدعى وجود عامل بدوام كامل لمدة أربعة أسابيع لإجراء اختبار مسبق. على الرغم من أن معظم المسوح لن تحتوي على العديد من الأسئلة إلا أنها تقدم مثلاً جيداً على أنواع الأسئلة التي يمكن استخدامها في برامج سبل عيش الشباب. من المهم كذلك إدراك أن العديد من النتائج قد لا يسهل قياسها (مثل سلوكيات المجازفة، الصحة العقلية، التمكين). تستخدم الاستبيانات المختلفة مداخلات مختلفة ويوصى باستخدام أدوات تم إعدادها في وقت سابق - بشكل مثالي المسوح التي تم التحقق من صدقها العلمي - للتوجيه.

تصميم الاستبيان واختباره

قبل أن يبدأ المسح في الميدان يجب إعداد الاستبيان ويتم هذا من خلال عملية تكرارية تستغرق عادة نحو شهر ولغاية شهرين.

الخطوة ١: التصميم

إن الاستبيان قائم على نتائج ومؤشرات تم إعدادها في وقت سابق. تعد اللغة المحلية واللهجات واللغة العامية المتداولة لدى الشباب مظاهر هامة يجب إدراجها وقد يلزم وجود مترجم للقيام بذلك أيضاً.

[معلومة إرشادية]

تتضمن الممارسات الجيدة لإجراء المسح على الشباب ما يلي:

- احصل على موافقة بعلم مسبق من كل من الشاب والوالد (أنظر القسم أدناه عن حماية مصالح المشاركين).
- استخدم لغة محلية مألوفة أو عامية، متى كان ملائماً.
- راعي قدرة الشاب على التركيز: اجعل المسوح قصيرة ومشوقة.
- استخدم أسئلة معقدة لتحسين جودة الاستجابات: قم بالإشارة إلى الماضي القريب لتنشيط الذاكرة والتذكر.
- كما هو الحال مع جميع المستجيبين: كن حذراً بشأن التوقيت وصياغة الأسئلة الحساسة.
- للمساعدة في العثور على الشباب لاحقاً، اجمع معلومات كثيرة عن العائلة والأصدقاء ومن يمكن الاتصال بهم في الأحياء المجاورة.
- إذا لزم معلومات عن الأسرة قم بتضمين وحدة مسح منفصلة تستهدف الآباء أو الأوصياء.

إذا كانت هناك مواضيع حساسة متضمنة في الاستبيان مثل طرح أسئلة عن الصحة العقلية أو العنف يجب صياغة الأسئلة بشكل فيه مراعاة للآخرين ويتمشى مع المعايير والعادات المحلية. تحتوي المسودة الأولى عادة على أسئلة سيتم في النهاية حذفها أو تغييرها.

الخطوة ٢: المراجعة الداخلية

بمجرد صياغة الاستبيان يجب أن يقوم أفراد الفريق الآخرين وأصحاب المصالح مثل مدير البرنامج ومسؤول المتابعة والتقييم والباحث الرئيسي وفريق العمل الميداني بمراجعته للتأكيد على أن الاستبيان يجمع كافة المعلومات اللازمة.

الخطوة ٣: الإختبار

يتم بعد ذلك أخذ مسودة الاستبيان إلى الميدان. كثيراً ما يتم إغفال هذه الخطوة برغم أهميتها من أجل إصدار تقييم جيد. إن الإختبار الميداني مهم للتأكد على أن طول الاستبيان وشكله وعباراته جميعها ملائمة ولضمان أن الاستبيان سيصدر نتائج متسقة وموثوق بها. يجب اختبار الاستبيان على مجموعة مختارة من الأفراد المشابهين للأشخاص الذين سيكونون في البرنامج ولكنهم لن يكونوا في العينة النهائية، وسيضمن هذا أن الأشخاص الذين سيتلقون الاستبيان النهائي غير متأثرين بتعرضهم بالفعل للأسئلة. من المهم كذلك إجراء اختبار مسبق للإجراءات التي ستستخدم في تحديد موقع الأشخاص الذين ستم المقابلة معهم لضمان إمكانية سهولة إيجادهم.

الخطوة ٤: المراجعة

تتم مراجعة مسودة الاستبيان لمعالجة القضايا التي تنشأ في الميدان ويمكن، إذا لزم الأمر، إعادة الخطوات حتى يتم حل جميع القضايا.

تدريب فريق العمل الميداني

عندما يكون الاستبيان معداً يجب تدريب فريق العمل الميداني على إدارته. يجب أن يعد خبير المسح أو مؤسسة لجمع البيانات دليلاً ليتم استخدامه كأداة تدريب ودليلاً إرشادياً للأشخاص الذين يجرون المقابلات. بحد أدنى، يجب أن يناقش الدليل أهداف المسح وإجراءاته متضمنة إجراءات التعامل مع المصاعب في الميدان. يجب شرح كل سؤال مسحي حتى يفهم الأشخاص الذين يجرون المقابلات الأساس المنطقي لتضمين السؤال في المسح. بالإضافة إلى ذلك، يجب أن يوفر الدليل للأشخاص الذين يجرون المقابلات إرشادات محددة عن كيفية طرح كل سؤال والحصول على معلومات ينتفع بها. يجب أن يراجع الباحث الرئيسي ومدير البرنامج الدليل. يقدم مربع ٧-٤ نموذج لموجز دليل مسح.

[مصدر إلكتروني]

أدلة التدريب على جمع البيانات
عنوان الموقع الإلكتروني
<http://www.iyfn.net/org/gpye-m&e-resource12>

١. أهداف المسح
٢. مهمات وأدوار وتوقعات الأشخاص الذين يجرون المقابلات والمشرفين وموظفي المسح الآخرين
٣. إجراءات مراجعة دقة البيانات
٤. مسح تفصيلي وإجراءات المقابلة (متضمنة إجراءات تحديد المستجيبين ومواقعهم والاتصال بهم بالإضافة إلى معلومات عن سلوكيات القائمين بالمسح من حيث السرية والموضوعية وسرعة المقابلة والانحياز والتعميق في الأسئلة)
٥. إرشادات عامة عن ملء الاستبيان والترميز
٦. شروحات بسيطة لكل سؤال
٧. إرشادات عن انتهاء المسح ومراجعته وتوجيه الشكر للمستجيبين
٨. إرشادات حول كيفية ملء التقرير الميداني وإخطار المشرفين بأية صعوبات تمت مواجهتها

[معلومة إرشادية]

يجب أخذ المعايير الحضارية والعادات المحلية بعين الاعتبار عند توظيف وتعيين مجري المقابلات. فعلى سبيل المثال، يكون من الجيد عادة استخدام موظفة تعداد عند إجراء مقابلة مع مستجيبات إناث، خاصة عند طرح أسئلة حساسة. ففي حال لم يشعر المستجيبون (أو الأوصياء عليهم) بالراحة مع موظف المسح، تصبح عدم مشاركتهم في المسح أكثر احتمالاً، وإذا شاركوا، عندها ستكون المعلومات المقدمة غير كاملة وغير دقيقة وبالتالي لا يمكن الاعتماد عليها.

يمكن أن يستغرق الأشخاص الذين يجرون مقابلات التدريب بضعة أيام أو ما يزيد عن الأسبوع، حسب مدى تعقيد المسح. يجب أن يبدأ التدريب بالإطلاع على المسح بأكمله سؤال بسؤال ثم يكون على كل شخص يجري المقابلة ممارسته على شخص آخر من يجرون المقابلة. يجب تشجيع الأشخاص الذين يجرون المقابلات على طرح الأسئلة أثناء هذه العملية لضمان فهم كل شخص لكل سؤال من الأسئلة. يجب أن تستمر هذه العملية حتى يألف جميع الأشخاص الذين يجرون المقابلات جميع الأسئلة. بعد انتهاء التدريب، يجب أخذ الأشخاص الذين يجرون المقابلات لموقع يمكنهم فيه إجراء الاستبيان على خمسة أشخاص على الأقل ممن يتشابهون مع عينة المستجيبين. يعد تدريب مجري المقابلة عملية تدريبية ومقابلة عمل في ذات الوقت. قم بدعوة ما نسبته عشرين بالمائة زيادة عن العدد المتوقع احتياجه للتدريب من مجري المقابلات ولا تقبل إلا الأفضل. إذا تم التعاقد مع مؤسسة للدراسات المسحية فستكون هي المسؤولة عن التدريب. كثيراً ما تعد فكرة وجود شخص من البرنامج لحضور بدايات الأيام الأولى من التدريب للإجابة عن التساؤلات التي تبرز فكرة جيدة. وهذه هي الفرصة الأخيرة للتخلص من أية أخطاء في الاستبيان.

حماية مصالح المشاركين

يمكن أن ينشئ البحث الذي يشارك فيه البشر معضلة في بعض الأحيان. فعندما تكون الغاية من بحثنا توليد معارف جديدة لفائدة برنامج محدد أو مجال بأكمله، على سبيل المثال عن طريق قياس أثر تدخل متعلق بسبل عيش الشباب، قد نميل إلى التفكير بنتائج تدخلنا على أنها أكثر أهمية من حماية الأفراد المشاركين في البحث. من الواضح أنه لا يجب أن نستخدم الشباب على أنهم مجرد وسيلة لغاية. ويوجد إجراءات معمول بها لمساعدتنا في تحديد قدرة تقييمنا على حماية المشاركين. بشكل أساسي، هناك ثلاثة مبادئ أساسية تحمي مصالح المشاركين في البحث (١)، صفحات ١٧-٢٠):

- احترام الأشخاص. يشير هذا المبدأ إلى ضمان فهم المشاركين المحتملين المخاطر والفوائد المحتملة للمشاركة في التقييم. ويعني هذا عملياً أنه يجب وضع خطة لضمان موافقة عن علم مسبق، والرغبة الصريحة للمشاركين الشباب في البحث في الإجابة على أسئلة المسح في ضوء فهمهم الواضح لطبيعة المسح.
- المنفعة. يشير هذا المبدأ إلى عدم إلحاق الأذى والزيادة القصوى لفوائد البحث المحتملة.
- العدالة. يتطلب المبدأ معاملة الأفراد والجماعات بعدل ومساواة فيما يتعلق بتحمل أعباء البحث وتلقي فوائده.

من أجل ضمان أعلى المعايير الأخلاقية في تقييم ما، سيطلب من العديد من الباحثين تقديم خطة تقييم الأثر الخاصة بهم للمراجعة من قبل مجلس مراجعة مؤسسية (IRB) في بلد المانح، والبلد المضيف، أو كليهما. ويفرض القانون هذه المراجعات على أي شخص يشارك في بحث تدعمه حكومة الولايات المتحدة والعديد من الحكومات الأخرى بالإضافة إلى أغلب الجامعات في جميع أنحاء العالم. وحتى لو لم تكن المراجعات مطلوبة من الناحية القانونية، فإن إجراء مراجعة الأخلاقيات هي فكرة جيدة بالنسبة لأي شخص يعمل مع المشاركين. وبشكل مثالي، يقوم مجلس المراجعة المؤسسية بمراجعة المسح قبل اختباره، ولكنه يقوم بهذه المراجعة بالتأكيد قبل تنفيذ المسح النهائي على نطاق واسع. ويمكن العثور على مجالس المراجعة المؤسسية في أية جامعة في الولايات المتحدة (وهذا الخيار الأفضل عند العمل مع باحث موجود في الولايات المتحدة) أو عبر مجلس محلي لمراجعة الأخلاقيات. وتقوم أيضاً معاهد أخرى مثل معاهد الصحة الوطنية الأمريكية أو الابتكارات لمكافحة الفقر بإجراء مراجعات الأخلاقيات عند الطلب. يُظهر مربع ٧-٥ نموذج استمارة مجلس مراجعة مؤسسية (IRB).

مربع ٧-٥ نموذج استمارة مجلس المراجعة المؤسسية (IRB).

عنوان الدراسة: _____

البلد والموقع: _____

التاريخ المتوقع للبدء والانهاء: _____

المحقق / المحققون، بما في ذلك اسم ومنصب وقسم ومؤسسة كل منهم: _____

١. غرض / خلفية / أهمية الدراسة، بما في ذلك سبب اعتبار الدراسة ذات قيمة.
 ٢. تصميم الدراسة، بما في ذلك كيفية تحديد مجموعات العلاج والمقارنة وتوقيت البرنامج. قم بوصف جميع التدابير التي يجب جمعها.
 ٣. قم بوصف المشاركين في الدراسة وإذا كان أي منهم من فئة السكانية الضعيفة، مع الإشارة إلى وجود أي تعويض للمشاركين.
 ٤. قم بوصف عملية الموافقة عن علم مسبق.
 ٥. هل يوجد أية مخاطر أو فوائد محتملة للدراسة؟
 ٦. كيف سيتم الحفاظ على السرية؟
 ٧. متفرقات: مذكرة تفاهم أو كتاب دعم من المنظمات الشركاء، والمسوح، ونماذج الموافقة، وشهادة تدريب المشاركين (بذخ أو ما يعادلها) لجميع موظفي البحث.
- المصدر: مقتبس من منظمة ابتكارات لمكافحة الفقر (٢٠١٠).

[تعريف]

إن مجلس المراجعة المؤسسية والذي يُعرف أيضاً بـ لجنة الأخلاقيات المستقلة، هو لجنة تم تعيينها رسمياً للموافقة على البحوث التي يوجد فيها مشاركون ومتابعة هذه البحوث ومراجعتها، وذلك لغرض حماية حقوق ورفاه هؤلاء الأفراد.

تشير الموافقة عن علم مسبق إلى الاستعداد الصريح لشخص (وعند الضرورة لوالده أو الوصي عليه) للمشاركة في البحث، ويفضل أن تكون هذه الموافقة كتابية. تتطلب الموافقة على علم مسبق كامل المعلومات حول جميع مقومات البحث التي يمكن أن تؤثر على استعداد الشاب للمشاركة.

عندما لا تمتلك منظمتك مجلس مراجعة مؤسسية موافق عليه

تمتلك تقريباً جميع المؤسسات الأكاديمية مجالس مراجعة مؤسسية، كما هو الحال في عدد من الوكالات المانحة والمنظمات غير الحكومية الدولية. وفي حال كنت تعمل بالشراكة مع واحدة من هذه الوكالات، قد يطلب منك أو يتم تشجيعك على إتباع إجراءاتهم من أجل الحصول على موافقة مجلس المراجعة المؤسسية. وفي حال كنت تعمل بشكل مستقل أو لا تستطيع الوصول لمجلس المراجعة المؤسسية الخاص بالشريك، تقدم العديد من الجامعات والمؤسسات الأخرى خدمات مراجعة الأخلاقيات. حيث يحتفظ مكتب حماية الأشخاص المشاركين في الأبحاث التابع لوزارة الصحة والخدمات البشرية الأمريكية بقاعدة بيانات قابلة للبحث تحتوي أكثر من ٨,٠٠٠ مجلس مراجعة مؤسسية حول العالم، من أفغانستان إلى زيمبابوي (انظر <http://ohrp.cit.nih.gov>). وبالإضافة إلى ما سبق، تُوفّر العديد من الوكالات المستقلة مراجعات للأخلاقيات، وتكون عادةً مجانية. للمزيد من المعلومات، راجع رابطة اعتماد برامج حماية المشاركين في الأبحاث (<http://www.aahrpp.org/www.aspx>)، واتحاد مجالس المراجعة المستقلة (<http://www.consortiumofirb.org/>).

عندما لا يوجد ما يكفي من الوقت لإجراء عملية الموافقة الكاملة لمجلس المراجعة المؤسسية

أولاً، قم بإعادة تقييم احتمال الحصول على مراجعة ضمن الوقت المتاح. إن برنامجك يتدخل بحياة الشباب وعائلاتهم، وتقع على عاتقك مسؤولية ضمان حماية المشاركين معك من الأذى بأفضل طريقة ممكنة. ولكن يمكن أن تستغرق موافقات مجلس المراجعة المؤسسية عدة أشهر، وقد تتعرض للضغط لبدء التنفيذ. إذا وُجد بعد التحليل الدقيق أنه بالفعل لا يوجد إمكانية للحصول على ترخيص مجلس المراجعة المؤسسية في الوقت المناسب، عندها وكحد أدنى، يجب أن يكون جميع أعضاء فريق التقييم قد تلقوا التدريب على حماية المشاركين في البرامج والأبحاث. وتوفر معاهد الصحة الوطنية (NIH) تدريباً مجانياً على شبكة الإنترنت (باللغتين الإنجليزية والإسبانية). للحصول على المزيد من المعلومات، انظر: <http://grants.nih.gov/grants/policy/hs/index.htm>

في الوقت الذي يُعتبر فيه احترام المعايير الأخلاقية أمراً جوهرياً في جميع مشاريع الأبحاث والتقييمات، قد تنشأ مسائل خاصة عند العمل مع الشباب، وتتطلب هذه المسائل اهتماماً إضافياً (أنظر جدول ٧-٣). حيث تجعل هذه المسائل من مشاركة مجلس مراجعة مؤسسية أمراً أكثر أهمية مما هو عليه في تقييمات أخرى، كما تتطلب تلقي الباحثين والاستشاريين المشاركين في التقييم تدريباً محدداً حول تطور الطفل والشباب وذلك قبل بدء التقييم. وبالإضافة إلى ما سبق، يجب تطوير بروتوكولات واضحة من أجل تحديد ماهية المعلومات التي سيتم جمعها وكيف سيتم استخدامها من أجل الحفاظ على أعلى المعايير الأخلاقية وحماية المشاركين.

المسائل	لماذا هذا مهم	ما الذي يجب فعله
معلومات عن مخاطر وفوائد المشاركة	قد يكون للشباب قدرة مختلفة عن الراشدين في إجراء تقييم دقيق للفوائد والمخاطر المرتبطة بالمشاركة في برنامج محدد أو مبادرة أبحاث محددة. وقد يكونون أيضاً أكثر استعداداً للمخاطرة بشكل عام، مما يجعلهم أكثر عرضة للعواقب السلبية المحتملة للمشاركة.	- توقع العواقب المحتملة بالنسبة للأطفال والشباب المشاركين. ولا تتابع إلا إذا كان بالإمكان منع العواقب الضارة المحتملة أو التخفيف منها. - أعط المشاركين من الشباب تفسيراً حول الهدف المقترح للبحث والإجراءات بلغة وبشكل مناسبين لعمرهم، ونصحبهم، وخبرتهم، ووضعهم. - وفر نقاشاً صريحاً حول أية مضايقات أو مخاطر قد يختبرها الشاب في حال وافق على المشاركة في البرنامج أو التقييم. - قل بصراحة أنه لا يوجد إلزام بالمشاركة في الدراسة وأن القرار بالمشاركة في الدراسة لن يكون له أي أثر على أهلية المشاركة في البرنامج. - لا تضع توقعات غير واقعية حول فوائد أو مكافآت المشاركة. وفي حال وجودها، قم فقط بتوفير مكافآت أو حوافز متواضعة للمشاركة بحيث تتماشى مع المعايير المعيشية المحلية.
الموافقة	قد لا يكون الشباب قد بلغوا سن الرشد القانوني بعد، ولهذا يجب طلب الموافقة من والديهم أو الأوصياء عليهم قبل إشراك الشباب أنفسهم. وعلاوة على ذلك، قد يكون من الصعب الحصول على رأي الشباب الصادق لأنه يتم عادة تنشئتهم اجتماعياً على الامتثال لآراء الكبار، بغض النظر ما إذا كانوا يوافقون على هذه الآراء أم لا.	- حدد عمر الأغلبية في البلد وقم بالاستشارة المحلية من أجل تحديد الجهة التي يجب أن تعطي الإذن للعمل مع الشباب (الوالدين، المدرسين، السلطات المحلية، القادة الاجتماعيين، الخ). - عند العمل مع القاصرين، اسعى دائماً للحصول على موافقة على علم مسبق من الوالدين أو الأوصياء. - إذا سمح عمر ونضج ووضع المشاركين الشباب، احصل أيضاً على موافقة عن علم مسبق من الشباب بالإضافة إلى موافقة والديهم.
جمع البيانات	إن جمع المعلومات حول المواضيع الحساسة (مثل استخدام العقاقير، والنشاط الجنسي، والتورط في الجرائم) أو التجارب المؤلمة (الإساءة، وفقدان الوالدين، والحرمان) هو أمر أكثر دقة عند التعامل مع الأطفال والشباب مقارنة بالبالغين. ويجب حماية مواطن ضعفهم العاطفية والجسدية.	- حاول قبل إجراء مقابلة مع الشباب جمع أكبر قدر من المعلومات من مصادر بديلة غير مباشرة (الراشدين، السجلات الإدارية، الخ). - قم بالاستشارة المحلية وصمم الاستبيانات، وأدلة مجموعة النقاش البؤرية/المركزية، والمواد الأخرى حسب خصائص المجموعة المستهدفة المحددة (مثل: احرص على أن تكون أدوات المسح مناسبة للعمر ومفهومة). - عند الضرورة، اعترف أنه يمكن أن تكون الأسئلة حساسة، وتوقع مخاوف الأهل والمشاركين وعالجها. - قل بصراحة أنه يمكن للمشارك الشاب رفض الإجابة على أي سؤال أو على جميع الأسئلة، وأن هذا لن يكون له أي أثر على الأهلية للمشاركة في البرنامج. ويمكن تكرار هذا النفي قبل طرح الأسئلة الحساسة.
السرية والحماية	إن حماية الخصوصية أمر حساس دائماً، وهو أكثر حساسية عند التعامل مع مستجيبين شباب ومواضيع حساسة. وبالنظر إلى مشاركة الوالدين أو الأوصياء الآخرين خلال عملية الموافقة ولعبهم دور الممثلين القانونيين، قد ينشأ عند العمل مع راشدين تناوب بين السرية والواجب الأخلاقي بحماية سلامة المستجيبين. فعلى سبيل المثال، قد يقوض حضور الوالدان للمقابلة خصوصية الشاب. وفي نفس الوقت، قد يكون هناك مسؤولية بإبلاغ الأوصياء في حال كان الشاب عرضة للآذى.	- اعمل دائماً على ضمان خصوصية وسرية الأجوبة من الوالدين والمشاركين الشباب، وهو أمر سيؤدي أيضاً إلى تعزيز موثوقية المعلومات المعطاة. - إياك أن تقم بنشر معلومات حول المستجيب دون الموافقة الصريحة من المستجيب والوالد. - خطط لكيفية التدخل في حال أعطى المستجيب معلومات توحى أنه نفسه أو أشخاص آخرين عرضة للآذى (من إساءة منزلية، وإهمال، وجريمة، وعنق)، أو أنه بحاجة لخدمات طبية أو قانونية أو خدمات أخرى. - في بداية كل مقابلة وبغض النظر عن الظروف الواضحة للمستجيب، قم بإبلاغ جميع المشاركين بالمصادر المتوفرة للرجوع إليها.

من أجل إجراء مسح لبرنامج التدريب المهني مي برير إمبليو والذي يستهدف الشباب الحضري في الهندوراس، تعاقد البنك الدولي مع مركز أبحاث الرأي الوطني (NORC) في جامعة شيكاغو لتهيئة تصميم الاستبيان وإدارة عملية جمع البيانات. وعلى الرغم من أنه لا يوجد في الهندوراس أية متطلبات قانونية للتعامل مع بيانات المسح الحساسة التي تنطوي على أشخاص مشاركون، إلا أن الشروط المرجعية للتقييم تطلبت موافقة مجلس مراجعة مؤسسية أمريكي من أجل تصميم البحث وخطة جمع البيانات، بالإضافة إلى إجراءات حماية البيانات التي تطابق المعايير الدولية. وبناء على هذا قدم مركز أبحاث الرأي الوطني جميع بروتوكولات واستبيانات البحث لمجلس المراجعة المؤسسية التابع لجامعته من أجل الموافقة قبل بدء العمل الميداني.

نظراً لطبيعة البحث، تم اختبار مجري المقابلات والمشرفين الميدانيين بخصوص خبرتهم في المسوح المتعلقة بالشباب. وخلال عملية تسجيل البرنامج، تم إبلاغ المتقدمين أنه سيطلب منهم المشاركة في مسح طوعي وأن قرارهم بالمشاركة في المسح لن يؤثر بأي شكل من الأشكال على اختيارهم للبرامج التدريبية. ونظراً لأن السن القانوني لمنح الموافقة في الهندوراس هو ١٨ سنة، طلب فريق جمع البيانات الموافقة المكتوبة من المستجيبين ممن يبلغون ١٧ سنة أو أصغر، وموافقة شفوية أو مكتوبة من والد أو وصي القاصر من أجل التسجيل في البرنامج، بالإضافة إلى موافقة منفصلة من القاصر والوصي من أجل المشاركة في مسح التقييم.

ولضمان السرية، تم الفصل الدقيق للمعلومات الشخصية عن نماذج المقابلة، واحتوت النماذج على مُعرّف رقمي فقط. وبهذا، كانت معلومات التسجيل الشخصية (الأسماء، والعناوين، إلخ) متوفرة حصرياً للمنظمة المنفذة (وزارة العمل والضمان الاجتماعي) لغرض الاتصال بالشباب الذين سجلوا، بينما تم تسليم بيانات الإجابة (بدون المعلومات الشخصية) للبنك الدولي من أجل التحليل فقط.

المصدر: مركز أبحاث الرأي الوطني (٢٠٠٧).

إجراء المسح الأولي والتحليل

يمثل المسح الأولي أول البيانات المجموعة حول مجموعات العلاج والمقارنة. وكما تم نقاشه سابقاً، فإن المسح الأولي ليس دائماً ضرورياً لجميع البرامج وأساليب تقييم الأثر. ولكن جمع البيانات الأساسية أمر مرغوب للغاية لأنه يوفر تقييماً مبكراً حول فعالية تصميم تقييم الأثر المختار من الناحية العملية، وفي نفس الوقت يوفر معلومات مفيدة حول خصائص المستفيد والتي يمكن أن تثرى البرنامج.

وسبب آخر لإجراء مسح أولي هو أنه يمكن أن يساعد في تحديد المشاركين فيما بعد. ويجب دائماً أن يتضمن المسح الأولي في حال إجرائه قائمة بمعلومات الاتصال من الشخص الذي أجري المسح معه، ومن الأصدقاء والعائلة الذين يمكن الاتصال بهم خلال مسح المتابعة.

التوقيت

يجب جمع البيانات الأولية قبل بدء البرنامج بقليل. ولو كان سيتم إجراء الجمع بعد بدء البرنامج، يمكن أن يكون البرنامج قد أثر بالفعل على الخصائص التي سيتم قياسها. وإذا تم إجراء المسح الأولي قبل البرنامج بوقت طويل، قد لا تعكس المعلومات المجموعة وضع المشاركين في بداية التدخل بدقة. إذا كنا نقوم بتقييم متوقع، يجب تعيين الأفراد في مجموعة العلاج والمقارنة قبل بدء البرنامج. ولكن لا يجب إبلاغ المشاركين في المسح بقرار التعيين هذا إلا بعد جمع البيانات الأولية.

[معلومة إرشادية]

للحصول على توجيهات مفصلة حول المناهج الأخلاقية للأبحاث المتعلقة بالأطفال والشباب، يرجى مراجعة:

Society for Research in Child Development. 2007. *Ethical Standards for Research with Children*. Available at http://www.srcd.org/index.php?option=com_content&task=view&id=68

Schenk, K. and Williamson, J. 2005. *Ethical Approaches to Gathering Information from Children and Adolescents in International Settings: Guidelines and Resources*. Washington, DC: Population Council. Available at <http://www.popcouncil.org/pdfs/horizons/childrenethics.pdf>

الإشراف على جمع البيانات

يعتبر تأكيد الجودة المفتاح لضمان أن البيانات التي تم جمعها ذات جودة مرتفعة. أولاً، من الهام إجراء اختبار الصدق لضمان أن القائمين على إجراء المقابلات قد استوفوا معايير وظيفتهم وأنهم استوفوا العدد المستهدف من عمليات المسح لكل يوم. فمن المعتاد تكوين فريق مستقل لمراجعة من ١٠-١٥ بالمائة من عمليات المسح ليثبت أن المستجيبين موجودون بالفعل وأنه تم جمع البيانات بشكل دقيق. وقد تساعد الحوافز على ضمان حفاظ القائمين على إجراء المقابلات على موقف إيجابي في وظيفة صعبة كهذه. وبالإضافة إلى الأجور، غالباً ما يحصل القائمون على إجراء المقابلات على بدل يومي يغطي مصاريف الوجبات والسكن أثناء سفرهم، بالإضافة إلى حوافز أخرى. ثانياً، يجب اتخاذ الخطوات اللازمة لحماية البيانات التي تم جمعها. فقد تفقد المعلومات إذا ضاعت الاستبيانات الكاملة أو سُرقت أو تعطلت الحواسيب الآلية. ولغرض عدم فقدان البيانات، يجب أن يتم جمع عمليات المسح في أسرع وقت ممكن من القائمين على إجراء المقابلات وتخزينها في مكان آمن. لذا، يجب دوماً نسخ نسخة احتياطية من بيانات الحاسب الآلي. وأخيراً، من الهام ضمان جودة إدخال البيانات. فقد يساعد استخدام أدوات إدخال البيانات الإلكترونية مثل الهواتف الجوالة أو المساعد الرقمي الشخصي، على عدم ارتكاب الأخطاء في إدخال البيانات، كما هو الحال لتدابير مراقبة الجودة المعيارية، مثل إدخال نفس البيانات مرتين.

التحليل ورفع التقارير

حالياً يتم جمع البيانات الأولية، يتعين على خبير التقييم الرئيسي ومساعد البحث استكمال التحليل والتقارير الأولية. وحيث لا يوجد هناك نتائج للبرنامج لرفع تقارير بشأنها بعد، سوف يحتوي التقرير الأولي على إحصائيات وصفية. فيجب مقارنة القيم المتوسطة للبيانات الديموغرافية لمجموعتي المعالجة والمقارنة وذلك لضمان ملاحظة أوجه التشابه بين المجموعتين والاختلافات ذات الدلالة الإحصائية. ويجب أيضاً عرض أية مشكلات قد تظهر مع جمع البيانات في التقرير الأولي (انظر مربع ٧-٦ لنموذج مخطط تفصيلي).

١. مقدمة
 - ١.١ وصف البرنامج والتقييم
 - ٢.١ فريق البحث
 - ٣.١ نظرة عامة على التقرير
٢. المعلومات الأساسية
 - ٢.١ الموقع والمكان
 - ٢.٢ الخلفية التاريخية
 - ٣.٢ الخلفية العلمية
 - ٤.٢ وصف البرنامج والشركاء المنفذين
٣. التدخل
 - ١.٣ اختيار المجموعة والمشاركين
 - ٢.٣ وصف التدخل
 - ٣.٣ مشاكل تتعلق بالتنفيذ
٤. تصميم تقييم الأثر
 - ١.٤ أهداف التدخل والنتائج المفترضة
 - ٢.٤ تصميم البحث والعشوائية
 - ٣.٤ مقاييس النتائج
 - ١.٣.٤ النتائج الأولية المرغوبة
 - ٢.٣.٤ النتائج الثانوية المرغوبة
 - ٣.٣.٤ النتائج العكسية
 - ٤.٣.٤ مقاييس أخرى موضع اهتمام
 - ٥.٣.٤ تدابير عدم التجانس
 - ٤.٤ المشكلات التي تم مواجهتها
 - ٥.٤ التدخل والمخطط الانسيابي والجدول الزمني للتقييم
٥. إدارة المسح الأولي
 - ١.٥ عمليات المسح الفردية والجماعية
 - ١.١.٥ تطوير المسح الأولي واختباره
 - ٢.١.٥ تعيين وتدريب موظف التعداد / شركة المسوح
 - ٣.١.٥ تنفيذ المسح الأولي
 - ٤.١.٥ المشكلات والمخاوف
 - ٢.٥ عمليات المسح الأخرى
٦. التحليل الأولي
 - ١.٦ السمات الأولية للمشاركين
 - ٢.٦ حسابات القوة واختبارات التوازن للبيانات الأولية
 - ٣.٦ الصديق الخارجي
 - ٤.٦ مشكلات جودة البيانات
٧. الاستنتاجات
 - ١.٧ المناقشات
 - ٢.٧ التفسير
 - ٣.٧ التعميم

ملحق

المصدر: مقتبس من يوز (٢٠١٠).

كما رأينا في بند ٦، يعتمد صدق كل أسلوب لتقييم الأثر على عدد من الافتراضات. ويمكن للتحليل الأولي أن يلعب دوراً هاماً في التحقق من هذه الافتراضات وذلك للتأكد بأن أسلوب التقييم الذي تم اختياره قابل للاستخدام، أو كيفية حل المشاكل في حال تم مواجهتها. ويقدم ملحق ٣ قائمة باختبارات التحقق والتزيف التي يمكن استخدامها في تقييم ما إذا الافتراضات الضمنية في تقييمنا المنشود حقيقية.

إجراء مسح وتحليل متابعة

عندما يعتمد أسلوب التقييم على جمع بيانات جديدة، سوف يقدم مسح المتابعة أو المسح الختامي البيانات التي طال انتظارها، والتي سوف تسمح لنا بتحليل ما إذا كان تدخلنا ناجحاً. فعندما يستند التقييم بالكامل على البيانات المتاحة، سوف يتم إجراء التحليل الخاص به أثناء هذه المرحلة.

التوقيت

سوف يحدد مدير البرنامج والمقيم الرئيسي مجتمعين توقيت مسح المتابعة. ولن تكون كل فائدة للبرنامج قابلة للملاحظة في الحال بعد التدخل، لذا يجب إجراء مسح متابعة بعد مرور وقت كاف للأثر حتى يتحقق. ويختلف الوقت طبقاً للبرنامج ويعتمد كثيراً على النتائج المحددة موضع الاهتمام. على سبيل المثال، قد يواجه الشباب المشاركون في برنامج تدريبي بعض المشاكل قصيرة الأجل فعلياً تتعلق بالمكاسب مقارنة بأقرانهم، وذلك لأنهم لا يستطيعون العمل أثناء فترة تدريبهم في الفصل الدراسي. ومع ذلك، إذا قدم برنامجنا للتدريب مهارات ذات صلة، فإننا نتوقع أن يحصلوا على دخل أعلى نسبياً خلال الأجل المتوسط والطويل. توقيت المتابعة سيكون مهم لتحديد التأثير الفعلي للتدخل. إذا رغبتنا في قياس النتائج سواء على الأجل القصير أو الأجل الطويل، فقد يلزمنا إجراء عدة مسح متابعة. وبالرغم من هذا سوف يزيد من تكلفة التقييم، إلا أنه سوف يحسن بشدة من قيمته. فيندر نسبياً استمرار تقييمات الأثر التي تتبع مجموعتي المعالجة والمقارنة على مدى أعوام عديدة وتكون تلك النتائج عادة أكثر طلباً وتقديراً. وإجراء أكثر من مسح متابعة واحد سوف يسمح لنا أيضاً بتحليل كيفية تغير نتائج البرنامج بمرور الوقت. ومع ذلك، إذا تأخر تنفيذ البرنامج، قد يبقى لنا وقتاً قليلاً للغاية بين نهاية البرنامج ونهاية الميزانية الخاصة بنا أو مدة المنحة لإجراء مسح المتابعة الذي سوف يكشف النتائج طويلة الأجل. وبناء عليه، من الهام إعادة ترتيب دورة التقييم إذا حدث تغييرات في الجدول الزمني للتنفيذ.

التتبع

أحد الاختلافات الرئيسية بين عمليات المسح الأولية والختامية هي مشكلة تتبع المستجيبين. فإذا لم يوجد الشباب موضوع المسح في مرحلة المتابعة، يمكن أن ينتج عن ذلك تحيزات خطيرة للغاية في التحليلات وتقليل قيمة ما تم التوصل إليه. على سبيل المثال، إذا قام المشاركون ذوو الأداء السيء بترك البرنامج، فمن المحتمل أن تكون نتائج التقييم مبالغاً في تقدير أثر البرنامج. وفي نفس الوقت، قد يترك البرنامج أكثر الشباب قدرة على الاستمرار. وحيث لا يمكننا بالتأكيد معرفة ما إذا سوف يؤدي معدل التناقص إلى تقدير مخفض أو مبالغ فيه للأثر، فمن الضروري تقليل التناقص لإجراء أي تقييم جيد. وعلى الرغم من أنه لا يمكن أبداً إيجاد ١٠٠ بالمائة من الأفراد الذين خضعوا للمسح من قبل، إلا أنه يجب بذل كافة الجهود لإيجاد أكبر عدد ممكن. ويتراوح معدل التناقص المقبول بصفة عامة ما بين ٥ إلى ١٥ بالمائة، مما يعني ضرورة إيجاد ٨٥ بالمائة من الشباب على الأقل في كل من مجموعتي المعالجة والمقارنة.

[معلومة إرشادية]

لغرض اعتبار نتائج التقييم النهائي جديرة بالثقة فيما بعد، من الجيد الاستعانة بالخبراء من الخارج في عملية المراجعة للتقرير الأولي والنهائي. وبالإضافة إلى ذلك، عند نشر التقرير الأولي، يمكن لموظفي البرنامج والتقييم خلق اهتمام عام في البحث الجاري وتقوية الملكية والحوار بين أصحاب المصالح من الداخل والخارج.

[معلومة إرشادية]

في كثير من الأحيان، يمكن تحديد مؤشرات بسيطة متسقة مع النتائج طويلة الأجل المتوقعة. على سبيل المثال، لن يتحقق دوماً أثر برنامج تدريب تنظيم المشروعات والبرامج الترويجية على إمكانية بدء مشروع تجاري لعدد من الأعوام (فالطلاب يتركون المدارس، ويحصلون على وظيفة ليكتسبوا خبرة ذات صلة، وفي النهاية يفكرون في بدء مشروع تجاري خاص بهم). وبقياس مؤشرات النتائج قصيرة الأجل ومتوسطة الأجل، مثل مهارات الأعمال، وتفضيل بدء مشروع تجاري كاختيار وظيفي، والخطوات الواقعية المتخذة لبدء مشروع تجاري، من الممكن الحصول على نتائج بسيطة للأثر دون الانتظار لعدة أعوام.

[تعريف]

التناقص يشير إلى معدل ترك المشاركين أو مستجبي المسح. وهذا يمثل مشكلة للتقييم لأن خصائص المشاركين الذين يقررون الترك تلقائياً ستكون مختلفة عن هؤلاء الذين يمكن إيجادهم، مما يشوه النتائج التي توصلنا إليها تبعاً لذلك. ويمكن أن يحدث معدل التناقص لعدة أسباب، منها على سبيل المثال، فقدان الاهتمام بالبرنامج، أو الهجرة، أو ببساطة عدم الرغبة في المشاركة في المسح.

في الشرق الأوسط، قامت الأمانة السورية بتوفير بطاقات شحن هواتف جواله لتحفيز الشباب على المشاركة في عمليات المسح. ولتوفير التكليف، طلبت الأمانة السورية من مشغلي الهواتف الجواله توفير هذه البطاقات كهبات عينية. وبالفعل، قدمت شركات الهواتف الجواله ١٠,٠٠٠ بطاقة بسعر ٢ دولار للبطاقة، بإجمالي مبلغ وقدره ٢٠,٠٠٠ دولار أمريكي. وبالنسبة لشركات الهواتف، كان ذلك دعاية جيدة بأقل تكلفة ممكنة.

في أوغندا، قام برنامج NUSAF بتعيين شركة لإجراء مسح تعقب مدته ١٠ دقائق للمستجيبين بعد المسح الأولي بعام واحد وقبل المسح الختامي بعام واحد. وقد طلب الاستبيان من المشاركين الذين كان من السهل إيجادهم، معلومات الاتصال المحدثة الخاصة بهم. وبالنسبة لهؤلاء الذين لم يكن من السهل إيجادهم، تم جمع المعلومات من الأصدقاء والأهل عن المكان المحتمل للشخص. ثم تم الاحتفاظ بهذه المعلومات للمسح الختامي لمساعدة الفرق في العثور على مجيبي المسح، بالإضافة إلى إعطاء إشارة للفرق حول مدى الصعوبة أو السهولة التي يمكن بها العثور على الأشخاص.

[معلومة إرشادية]

هناك طرق إضافية لتيسير التتبع، تتضمن ما يلي:

- طلب المشورة والمساعدة من القادة والمسؤولين والمقيمين المحليين. فقد يعرف المحليون أفضل طريقة لإيجاد شخص ما.
- إشراك القائمين على المسح من المكان موضع الدراسة لأنهم على دراية بالمنطقة والأعراف المحلية.
- إذا كان لا يزال من غير الممكن العثور على المشاركين، قم باختيار عينة عشوائية من هؤلاء الذين لم يتم العثور عليهم وذلك لإجراء بحث مكثف للغاية عنهم. وفي حالة الاختيار العشوائي، يمكن اعتبار هؤلاء الذين تم العثور عليهم في النهاية من خلال البحث الحديث، ممثلين للآخرين الذين لم يتم العثور عليهم.

إن تتبع الأشخاص، وبصفة خاصة الشباب كثيري التنقل، يمكن أن يكون صعباً. وهناك ثلاث طرق شائعة للحد من معدل التناقص، وهي كما يلي:

- **جمع معلومات اتصال جيدة أثناء المسح الأولي.** يجب أن يتضمن المسح الأولي أنواع عديدة من معلومات الاتصال (عنوان الشارع، عنوان البريد الإلكتروني، رقم الهاتف، إلى آخره) من المستجيب ومن الأصدقاء والأسرة أيضاً الذي يمكن أن يساعدوا في إيجاد الشباب لمسح المتابعة. كذلك، فإن استخدام قنوات التواصل الاجتماعي مثل الفيسبوك يمكن أن تساعد أيضاً على تتبع الشباب.
- **تحفيز الشباب في مجموعتي المعالجة والمقارنة حتى يتواجدوا عند إجراء عمليات المسح في المستقبل.** يمكن أن تتضمن الحوافز للمشاركة في عمليات مسح المتابعة، مبالغ بسيطة مقابل وقتهم أو يانصيب لمبالغ نقدية أو جوائز. ويمكن إخطار الشباب بهذه الحوافز من خلال وسائل الاتصال المرتبة مسبقاً (ربما عند بداية المسح)، أو من خلال وسائل الإعلام الجماهيرية مثل الإذاعة وإعلانات الصحف.
- **استخدام مسح تتبع.** في حالة التقييمات التي يوجد فيها فاصل زمني كبير بين البداية والختام مثل عامين أو أكثر، ولاسيما لتلك التقييمات التي لا تستخدم مسح أولي، يمكن استخدام مسح تتبع قصير وسريع للغاية لتقدير معدل التناقص المحتمل وجمع معلومات إضافية. وإذا كان البرنامج ذو ميزانية مقيّدة، يمكن النظر أيضاً إلى إجراء عمليات مسح متابعة عن طريق الهاتف للحصول على معلومات اتصال محدثة من مجيبي المسح، وتخصيص الزيارات الشخصية للشباب الذين لا يمكن الوصول إليهم عن طريق الهاتف.

تصميم مسح المتابعة وجمع البيانات

من المحتمل أن يرغب فريق البرنامج أو التقييم في إضافة بعض الأسئلة الإضافية على المسح الأصلي (انظر مربع ٧-٧)، والتي قد تتضمن أسئلة حول معدلات الحضور، وحالات الترك، والدوافع لكل منهما، لأن هذه المعلومات يمكن استخدامها لتقدير مدى الخدمة الذي قد تلقاها الأفراد بالفعل. وسوف يلزم تجريب الأسئلة الجديدة ومراجعتها كلما كان ذلك ضرورياً. وبصفة عامة، من الأفضل الحفاظ على أسئلة المتابعة وترتيب الأسئلة على نحو مشابه للمسح الأولي على قدر الإمكان لضمان تشابههم. وما لم يكن هناك مشكلة كبيرة بخصوص سؤال في المسح الأولي، فمن الأفضل أن تترك صياغة السؤال كما هي في عمليات مسح المتابعة. وسوف يحتاج دليل المسح أيضاً إلى تحديثه ليعكس أية تغييرات مختلفة عن المسح الأولي. وبصفة خاصة، يجب أن يتضمن بروتوكولات محددة لتتبع المشاركين في المسح.

مربع ٧-٧ أنواع الأسئلة الشائعة التي ستضاف إلى مسح المتابعة

- أسباب عدم المشاركة أو ترك البرنامج.
- معدل حضور المشاركين أو حجم المنافع والفوائد التي يتلقونها.
- مدى رضا المشاركين بالبرنامج.
- تقييم المشاركين لجودة البرنامج.
- تقييم المشاركين الذاتي للنواتج من البرنامج.

وأخيراً، سيحتاج القائمون على إجراء المقابلات لنفس مستوى التدريب والإشراف الذي تم تقديمه أثناء المسح الأولي لضمان أفضل جودة لجمع البيانات. وإذا كان من الممكن، قم باختيار أفضل القائمين على إجراء المقابلات من المسح الأولي لإجراء مسح المتابعة. ويجب استبدال أو تقديم برامج تدريب إضافية للقائمين على إجراء المقابلات ممن يرتكبون معدلات مرتفعة من الأخطاء أو ممن هم أقل اعتمادية.

التحليل النهائي وتقرير التقييم

بعد جمع بيانات المتابعة، يمكن إصدار تقرير تقييم الأثر النهائي، والذي يمثل المنتج النهائي للتقييم. وسوف يكرر التقرير النهائي كثير من المعلومات المبينة في المسح الأولي، وسوف يضيف معلومات تفصيلية عن المسح الختامي وتحليل البيانات النهائية.

سوف يستند التحليل إلى النتائج والمتغيرات المحددة سابقاً. وفي بعض الحالات النادرة، يمكن إجراء التحليل بإجراء مقارنة بسيطة للمتوسط بين مجموعتي المعالجة والمقارنة (عادة في حالة تصميمات اليانصيب). ومن الناحية العملية، سيستخدم في كثير من الأحيان شكل ما من أشكال تحليل الانحدار للتحكم في عدة متغيرات رئيسية التي قد تجعل النتائج متحيزة.

يقدم مربع ٧-٨ نموذجاً لموجز أقسام تقرير التقييم الذي يمكن إضافته للتحليل الأولي. جميع هذه المعلومات هامة لضمان أن الشخص الذي لم يشترك في التقييم يستطيع أن يفسر النتائج بشكل صحيح.

[تعريف]

في علم الإحصاء، يتضمن تحليل الانحدار أية أساليب لوضع نماذج لمتغيرات عديدة وتحليلها. في تقييم الأثر، يساعدنا تحليل الانحدار في فهم كيف تتغير القيمة النموذجية لمؤشرات النتائج بين مجموعة المعالجة أو المقارنة بينما يتم الإبقاء على خصائص المستفيدين ثابتة.

مربع ٧-٨ مثال للإضافات على التقرير الأساسي بعد الاستبيان النهائي

٧. إدارة المسح الختامي

١.٧ مسح ختامي للأفراد والمجموعات

١.١.٧ تطوير المسح الختامي واختباره

٢.١.٧ تعيين وتدريب شركة المسوح / مجري المقابلة

٣.١.٧ التعبئة وبروتوكولات التتبع

٤.١.٧ تنفيذ المسح الختامي

٢.٧ البروتوكولات النوعية

٣.٧ المشكلات والتأخيرات

٤.٧ قضايا جودة البيانات

٨. تحليل البيانات

١.٨ الأساليب الإحصائية المستخدمة

٢.٨ مستويات التحليل

٣.٨ ملخص النتائج

٤.٨ تحليلات مكملية

٩. الاستنتاجات

١.٩ مناقشات

٢.٩ التفسير

٣.٩ قابلية التعميم

٤.٩ توجيهات للأبحاث المستقبلية

الملحق

المصدر: مقتبس من يوز (٢٠١٠)

فهم عدم التجانس

قد لا يستفيد كافة المستفيدين من البرنامج من تدخلنا بنفس الطريقة، لذلك، تتمثل إحدى نقاط أهمية التقييم في فهم التباين في آثار البرنامج. على سبيل المثال، ترغب العديد من البرامج في معرفة من يحقق أداء أفضل في البرنامج؛ الصبية أم الفتيات أو الشباب الأصغر أم الأكبر أو أصحاب المستويات الأعلى أم الأدنى من التعليم أو الخبرة. بالإضافة إلى النظر إلى النوع أو العمر أو التعليم، قد نرغب كذلك في تقييم ما إذا كانت النتائج تختلف طبقاً للثروة المبدئية للمشاركين (قيمة أصول المشاركين) ورأس المال الاجتماعي (الإستفادة من التشبيك) أو السمات النفسية (التفاؤل وسلوكيات المجازفة وما شابه). إن فهم أي من المشاركين قد حقق الاستفادة القصوى وأيهم قد حقق الاستفادة الأدنى من برنامجنا من شأنه أن يساعدنا في تصميم التدخل أو استهدافه بشكل أفضل. (للحصول على مزيد من المعلومات عن قياس عدم التجانس، أنظر قياس مجموعة متنوعة من الآثار في بند ٨.)

[مصدر إلكتروني]

تقارير تقييم الأثر
عنوان الموقع الإلكتروني
<http://www.iyfnet.org/gpye-m&e-resource13>

[تعريف]

عدم تجانس الأثر يشير إلى الاختلافات في الأثر طبقاً لنوع المستفيد، أي كيف تستفيد المجموعات الفرعية المختلفة من تدخل ما إلى حد مختلف.

على سبيل المثال، إذا وجد تقييمنا أن برنامج تدريب سبل العيش أثره أكبر على الرجال، يمكن أن تركز التكرارات المستقبلية من البرنامج على الرجال بشكل أكبر لزيادة المردود النهائي للبرنامج. ومن جهة أخرى، وبناء على الأولويات، يمكننا اكتشاف طرق لحث النساء على الاشتراك بحيث يستفدن كذلك من البرنامج.

ومع ذلك، على النحو المبين في مربع ٧-٩، يجب تحديد عدم التجانس مسبقاً قبل أي تحليل ويجب الإبلاغ عن جميع النتائج ليس فقط النتائج التي وُجدت ذات دلالة إحصائية. نحن نرغب في تجنب استقصاء البيانات الأمر الذي يمثل مشكلة كبيرة عند تحليل عدم التجانس.

مربع ٧-٩ التنقيب عن البيانات

يعد التنقيب عن البيانات مشكلة كبيرة في علم الإحصاء. ويشكل معضلة بوجه خاص مع المسوح الطويلة للغاية التي تطرح العديد من الأسئلة، غالباً بطرق مختلفة.

فيما يتعلق بالتنقيب عن البيانات، يسعى الشخص إلى الحصول على النتائج التي تؤكد معتقدات خاصة عن برنامج ما وتتجاهل النتائج التي لا تؤكد هذه المعتقدات. على سبيل المثال، قد تعتقد مسؤولة البرنامج اعتقاداً قوياً أن برنامج التدريب له أثر إيجابي على الشباب. وبمجرد أن تحصل المسؤولة البيانات من التقييم، تجد أن هناك زيادة ذات دلالة إحصائية في الوقت المستغرق في العمل ولكن متوسط دخل الشباب ليس ذو دلالة إحصائية متزايدة. فالتقرير وجود الزيادة في الوقت المستغرق في العمل فقط وليس حقيقة عدم حدوث تغيير في الدخل هو نوع من أنواع التنقيب عن البيانات.

يمكن أن يحدث التنقيب عن البيانات بطريقتين. الأولى عندما نتجاهل الأدلة التي تتناقض مع معتقداتنا ونذكر فقط الأدلة التي تؤكد معتقداتنا أما الطريقة الثانية فهي القيم الإحصائية الشاذة. في علم الإحصاء، توجد دائماً فرصة أن يظهر متغير بأنه ذا دلالة. ولكن في الحقيقة، على الأقل ٥ بالمائة من الوقت، سيظهر أن شئ ما ذو دلالة في حين أنه في الحقيقة ليس كذلك. إذا جمع المقيم ١٠٠ معلومة ستوصف ٥ منها على الأقل بأنها ذات دلالة في حين أنها ليست كذلك. إذا بحث الباحث عن هذه المعلومات الخمسة وذكرها فقط في تقريره فستكون النتائج في الحقيقة غير صحيحة.

قد يجد تقييم ما أنه لا يوجد أثر ذو دلالة إحصائية للبرنامج. ولكن بالتعمق في حالات عدم التجانس، وبسبب العشوائية الإحصائية، سيد الباحثون أثر ما على مجموعة ما. لتجنب التنقيب عن البيانات، يجب أن نحدد جميع النتائج موضع الاهتمام قبل إجراء التحليل وأن نذكر جميع هذه النتائج دون إخفاق، متضمناً النتائج التي لم يتم العثور على أثر لها. بهذه الطريقة، يمكن فهم الصورة كاملة.

تفسير النتائج

جودة التنفيذ: تعتمد النتائج بشكل كبير على مدى جودة تنفيذ التدخل. يجب إذن أن يناقش تقرير التقييم النهائي جودة التنفيذ بالتفصيل. تعد المعرفة الجيدة عن الكيفية التي نُفذ بها البرنامج أمراً مهماً بالأخص عندما تُظهر نتائج التقييم أثراً محدوداً أو سلبياً. بما أنه يسمح لنا بالتفريق بين المشاكل المتعلقة بالتنفيذ والمشاكل المتعلقة بالتصميم. حتى نكون قادرين على تفسير نتائج التقييم بدقة، فمن الضروري تضمين تقييم الأثر في إطار من المتابعة القوي وتقييم العمليات وغيرها من الأدوات النوعية الأخرى.

درس كل من برون وزيا (٢٠١١) أثر برنامج شامل نحو الأمية المالية وأمية الأعمال على النتائج لأصحاب الأعمال الشباب في اقتصاد ناشئ ما بعد النزاع للبوستة والهرسك. على الرغم من أنهم لم يجدوا متوسط آثار متميز لبرنامج التدريب على أداء الأعمال، إلا أنهم وجدوا مستويات مرتفعة من عدم التجانس بين المشاركين. بوجه خاص، وُجد أن أصحاب الأعمال الشباب الذين يملكون معرفة مالية مسبقة قبل البرنامج أظهروا تحسينات في المبيعات بسبب برنامج التدريب. وكذلك كانت التأثيرات على الأرباح إيجابية بالنسبة لهذه المجموعة الفرعية. توحى النتائج بأن التدريب لا يجب أن يكون التدخل الوحيد لدعم أصحاب الأعمال الشباب وأن محتوى البرنامج الدراسي المحدد كان ملائماً لمجموعة محددة للغاية من أصحاب الأعمال الشباب ولكن ليس للجميع.

[معلومة إرشادية]

يُعد وجود بيانات جيدة عن الحضور من متابعة البرنامج أمر مفيد للغاية حيث إنه لا يخبرنا فقط عن عدد الشباب الذين التحقوا بل أيضا يخبرنا إلى أي حد استخدمت الخدمات المقدمة. وهذا يسمح بالتمييز بين المشاركين المنتظمين وغير المنتظمين وتحديد ما إذا ترك أحد البرنامج في وسطه (يحتمل أن يكون قد تم استبداله بشخص آخر). إذا لم يتم جمع هذه المعلومات وتحليلها فعلى الأرجح سيخس تقييم الأثر من قدر فاعلية البرنامج. تساعدنا هذه المعلومات كذلك على فهم تأثير الجرعات المختلفة: على سبيل المثال، الاختلاف في نتائج شخص ما تلقى ١٠٠ ساعة من التدريب مقابل شخص آخر تلقى ٥٠ ساعة فقط.

قابلية النتائج للتعميم: مثالياً، هناك صدق خارجي لتقييم الأثر الخاص بنا، مما يعني أنه يمكننا تعميم نتائجنا على سكان آخرين وظروف أخرى. يعتمد وجود هذه الحالة بشكل كبير على إستراتيجية اختيار العينات المختارة في التقييم. كلما كانت العينة أكثر تمثيلاً، كلما زادت ثقتنا في أن البرنامج يمكن أن ينجح كذلك مع مجموعة مستفيدين أكبر أو مجموعة مستفيدين مختلفة. إن لهذا آثاراً هامة من ناحية قابلية توسيع نطاق التدخل وتكراره. بوجه عام، من الصواب افتراض أن التغييرات عبر الزمن، والبيئات المختلفة، وآليات التطبيق المختلفة من موقع لآخر يكمن بها احتمال التأثير بشكل ملحوظ على أثر البرنامج في اتجاه أو الآخر. يجب إذن الحرص دائماً عند ترجمة دروس التقييم من برنامج لآخر وإدراك أن المتابعة والتقييم سيكونان دائماً ضروريان للتعلم المستمر وتحسين البرنامج.

نشر النتائج

بمجرد الحصول على نتائج تقييم الأثر تكون الخطوة الأخيرة وتمثل في نشر النتائج للعاملين بالبرنامج بالإضافة لمن قد يكونون مهتمين بالنتائج من خارج البرنامج.

النشر الداخلي

يوفر النشر الداخلي للتقييم أساساً للتعلم التنظيمي. إن مشاركة النتائج مع العاملين بالبرنامج وبقية المنظمة يلبي دافعاً من الدوافع الرئيسية لإجراء التقييم في المقام الأول: تعزيز إدارة البرنامج (أنظر بند ١). حتى ينتج الاهتمام والإحساس بالملكية، من الأفضل البدء فوراً في عملية النشر الداخلي بعد استكمال التقييم الأولي - على سبيل المثال، بالمشاركة في النتائج الأولية وعرضها ثم يجب بعد ذلك نشر نتائج التقييم على الموظفين التنفيذيين والقادة في المكاتب القطرية والمراكز الرئيسية متى كان هذا قابلاً للتطبيق. يمكن أن يتضمن التقرير مناقشة حول إمكانية تأثير النتائج على تصميم المبادرات المستقبلية أو الحالية.

النشر الخارجي

يجب أن يستهدف النشر كذلك أصحاب المصالح الخارجيين مثل السلطات المحلية والوزارات القومية والمنظمات غير الحكومية المحلية والدولية والجامعات (خاصة أقسام التنمية والاقتصاد والصحة العامة)، والمنظمات متعددة الأطراف (مثل الأمم المتحدة والبنك الدولي وبنوك التنمية الإقليمية) أو المانحين الثنائيين (مثل الوكالة الأمريكية للتنمية الدولية (USAID) والجمعية الألمانية للتعاون الدولي (GIZ) ووزارة التنمية الدولية (DFID). في الحقيقة، هناك طلب متزايد على نتائج تقييم الأثر بوجه عام خاصة في مجال سبل عيش الشباب حيث مازالت الأدلة الدقيقة عما ينجح وما لا ينجح نادرة. توجد طرق متعددة للوصول إلى الجمهور الخارجي وعادة ما تستخدم خطط النشر القنوات على شبكة الانترنت والقنوات المباشرة (أنظر مربع ٧-١٠). إن نتائج التقييم التي يتم مشاركتها بشكل واسع يمكن أن يكون لها تأثير متدفق في جميع أنحاء العالم.

النشر على شبكة الانترنت

- الموقع الإلكتروني للمنظمة
- النشرات
- بوابات المعرفة على شبكة الانترنت (تحميل التقرير والنتائج)
- سجل وظائف الشباب <http://www.youth-employment-inventory.org/>
- موقع مجموعة شبكة وظائف الشباب <http://yenclinic.groupsite.com/>
- إديس <http://www.eldis.org/>
- زونيا <http://zunia.org/>
- قواعد بيانات الأوراق البحثية
- أوراق مناقشة جمعية زيوليت الدولية
- <http://www.iza.org/en/webcontent/publications/papers>
- شبكة أبحاث العلوم الاجتماعية
- <http://papers.ssrn.com/sol3/DisplayAbstractSearch.cfm>
- أوراق ايكون <http://econpapers.repec.org/>
- المدونات ووسائل التواصل الاجتماعية

النشر المباشر

- المؤتمرات المتعلقة بمواضيع متخصصة
- المؤتمر العالمي للفرص الاقتصادية للشباب
- <http://www.youtheconomicopportunities.org>
-
- المؤتمر السنوي سيب SEEP
- <http://www.seepnetwork.org/Pages/conference.aspx>

العروض

- المنظمات الدولية (البنك الدولي، البنك الإسلامي للتنمية (IDB)، منظمة التعاون الاقتصادي والتنمية (OECD)، مكتب العمل الدولي (ILO)، اليونيسيف (UNICEF)، برنامج الأمم المتحدة الإنمائي (UNDP)، وغيرها).
- المانحون الثنائيون (مثل الوكالة الأمريكية للتنمية الدولية (USAID)، والجمعية الألمانية للتعاون الدولي (GIZ)، ووزارة التنمية الدولية (DFID)، الوكالة الفرنسية للتنمية (Afd)، وغيرها).
- الجامعات (المحلية والدولية)

منتجات تكميلية

ملخصات السياسات

تساعد ملخصات السياسات في توصيل النتائج إلى أصحاب المصالح الداخليين والخارجيين. إن ملخص السياسة (وهو لا يزيد غالباً عن أربع صفحات) يقدم النتائج الأساسية للتقييم على نموذج مكتوب بوضوح يتضمن رسوم بيانية وجداول ويطرح توصيات برامجية وسياسية.

[مصدر إلكتروني]

أمثلة على المنتجات التكميلية
عنوان الموقع الإلكتروني
<http://www.iyfnet.org/gpye-m&e-resource8>

أوراق العمل

قد يعمل خبير التقييم مع فريق البرنامج على كتابة أوراق العمل والمقالات لنشرها في الدوريات الأكاديمية ولتقديم نتائج البحث في الجامعات ومعاهد الأبحاث. يمكن عندئذ إصدار أوراق العمل ونشرها من خلال الجمعيات الأكاديمية حيث ينتمي الباحثون. حيث يعد الاقتباس منها في الأوراق الأكاديمية طريقة ممتازة لزيادة التعرف على البرنامج وخلق اهتمام بين المانحين.

كشف المشكلات وحلها

كما هو الحال مع أي برنامج أو تقييم، يكون من الشائع مواجهة مشكلات عند إجراء تقييم للأثر. تقدم القائمة التالية أمثلة على قضايا شائعة في الخطوات المختلفة لتقييم الأثر وكيف يتم تجنبها أو تخفيف حدتها.

الإعداد للتقييم

تقييم البرنامج الخطأ. يمكن إهدار الكثير من الأموال على تقييمات للأثر تكون منافعها ومساهماتها غير واضحة. نظراً لمحدودية الموارد، فمن المهم استهداف تقييمات الأثر في التدخلات الاستراتيجية وغير المختبرة والتي قد تخضع للتكرار والتوسيع.

الأهداف غير الواقعية. تعاني العديد من التدخلات من «انحراف الرسالة»، الذي وفقاً له يتغير الهدف من البرنامج بمرور الوقت. من الصعب وضع مؤشرات تقييم مفيدة في ظل هذه الظروف. بالمثل، تحديد أهداف غير واقعية من ناحية النتائج يؤدي على الأرجح إلى نتائج تقييم تُظهر عدم وجود أثر على هذه النتائج. من المهم أن نكون واقعيين عند تعريف النتائج المنشودة والأهداف التعليمية من التقييم.

التأثيرات الخارجية. حتى بعد تصميم تقييم محدد قد تعرقل العوامل السياسية المضي قدماً في استراتيجية التقييم المختارة. بشكل بديل، يمكن أن تُعجل العوامل الخارجية التنفيذ أو تؤخره، مما يؤثر على تقديم الخدمات والتقييم، مثلاً من خلال التدخل المتأخر أو غير المتسق أو تلوث مجموعتي المعالجة والمقارنة. وإحدى الطرق الممكنة لتخفيف تأثير طرف الثالث هي الموافقة الحازمة على خطة التنفيذ والتقييم (مثالاً عمل مذكرة تفاهم) ومراجعتها على نحو دوري.

تحديد إطار زمني وميزانية

التخطيط غير الواقعي: عند إعداد الجدول الزمني، تتمثل المخاطرة الرئيسية في التقليل من الفترة الزمنية والموارد المطلوبة لإجراء تقييم الأثر كما ينبغي. فمن الشائع التعرض إلى تأخيرات في تصميم وتنفيذ البرنامج، والتي تزيد تباعاً من الفترة الزمنية - وعلى الأرجح من تكلفة - التقييم. على سبيل المثال، قد ينتج عن التأخيرات عدم توافر العاملين والاستشاريين الرئيسيين لفترة أطول. ومن ثم فمن الضروري وضع ميزانية معتدلة والتطلع المستقبلي للعاملين.

إعداد فريق تقييم

التوظيف. إن توظيف فريق جيد لتقييم الأثر، بدءاً من تحرير الشروط المرجعية حتى تحديد خبراء وشركات مؤهلة، يمكن أن يمثل تحدياً. كما أن التقليل من الخبرة المطلوبة في المراحل المختلفة واستخدام أشخاص غير مناسبين يمكن أن يؤدي إلى تأخيرات واضحة وتجاوز للتكلفة، الأمر الذي يؤدي في النهاية إلى ضعف نتائج التقييم. من الضروري ضمان أن شروط كل دور من الأدوار قد تم تحديدها مقدماً واستيفائها من قبل الخبير أو الشركة المختصة. كما أن العمل مع المؤسسات القائمة (مثل الجامعات ومؤسسات الفكر) والتي يكون لديها سجل معروف في إجراء دراسات بحثية ذات جودة، يمكن أن يساعد في توفير الدعم على المستوى المحلي وضمان قبول النتائج النهائية على نطاق واسع.

تغيير العاملين. تستبدل الشركات التي تفوز بجهات اتصال للتقييم في بعض الأحيان موظفين أقل خبرة بالعاملين الرئيسيين. الأمر الذي يمكن تجنبه من خلال إدراج بنود تعاقدية تفرض عقوبات أو تتخذ إجراءات تصحيحية في هذا الشأن.

إدارة فريق المسح. تصبح إدارة الفريق الداخلي للمسح أكثر تعقيداً بشكل سريع. فعند جمع البيانات مع العاملين بالبرنامج، يجب التأكد من فهم احتياجات جميع العاملين وضمان وجود إشراف كاف واستخدام الإدارة على النحو المطلوب للتعامل مع الفريق.

إعداد خطة تقييم

القصور في البيانات القائمة. عند الاعتماد على بيانات ثانوية، يجب التأكد من توافرها وجودتها. قد لا تطرح المسوح القائمة الأسئلة ذات الصلة بتقييم محدد لدينا أو تخاطب مجتمع الدراسة موضع اهتمامنا، أو قد يكون لديها حجم عينة صغير للغاية لا يتلاءم مع قوة دراستنا. فمن المهم فهم القصور فهماً كاملاً قبل الالتزام باستخدام البيانات القائمة فقط.

الانفصال بين البرنامج والتقييم. إن التواصل والتنسيق غير الكافيين بين المنظمة التنفيذية والمقيم الأساسي قد يؤديان إلى اختيار تصميم للتقييم غير قابل للتنفيذ من الناحية العملية. إن الاحتفاظ بمشاركة العاملين الأساسيين بالبرنامج في إعداد خطة التقييم يمكنه أن يضمن تناسب التقييم مع السياق التشغيلي. إلا إنه في حالة وقوع انفصال وتم اكتشافه مبكراً، يكون الحل الأفضل هو البحث عن أسلوب للتقييم أكثر واقعية.

تحيز الانتقاء. إن تحديد العينة بدقة والاختيار العشوائي للمشاركين في الدراسة هو الطريق الأبسط والأكثر صموداً للتخلص من التحيز في الانتقاء. إذا ظهر تحيز الانتقاء في البيانات، يكون من أحد الحلول المنقوصة مقارنة النتائج المجموعة المعالجة مع عينة مطابقة تم سحبها من مجموعة بيانات مختلفة.

إعداد وتجريب أداة المسح

هوس القياس. هو استهداف العديد من النتائج ومن ثم اشتغال أدوات المسح على العديد من الأسئلة والتي تؤدي في الغالب إلى رفع تكلفة المسح بما يفوق ميزانية المسح. إن طرح العديد من الأسئلة يزيد من الأعباء التي تقع على عاتق مشاركي المسح وقد يؤدي إلى انخفاض معدل الإجابات وجودتها. اقتطاع الأسئلة المتعلقة بالنتائج غير المباشرة طريقة مناسبة للحد من هذه القضية.

الاختبار غير الكافي. إن الخطوة التي يتم غالباً تجاوزها لصالح توفير الوقت هي تجريب أدوات التقييم. يعتبر التجريب خطوة حاسمة في سير العملية والتي لا يمكن استبعادها، ولا سيما أن إجراء مسح على الشباب يفرض تحديات إضافية ربما لا يتم فهمها على الفور. إذا لم يتم التأكد من صحة أداة التقييم، فقد تأتي النتائج غير دقيقة أو غير مكتملة أو مضللة. لذا يتعين استغراق الوقت اللازم خلال مرحلة الاختبار الميداني للمسح لضمان أن المعلومات التي تم جمعها جاءت على أعلى قدر من الجودة.

خفض معايير الأخلاق. قد يؤدي إجراء مسح دون الحصول على موافقة من مجلس المراجعة المؤسسية أو من اللجنة المحلية للأخلاقيات إلى صد هائل من أصحاب المصالح وقد تؤدي كذلك إلى فقدان أهلية التقييم برمته. يُعد تدريب القائمين على الأخلاقيات الأساسية لإجراء التقييم الحد الأدنى من الشروط المطلوبة.

إجراء مسح وتحليل أولي

إيجاد مستجيبين. قد يصعب تحديد مواقع الشباب لإجراء المسح، وفي هذه الحالة، يُنصح بطلب مشاركة العاملين بالبرنامج على المستوى المحلي وغيرهم من أصحاب المصالح في البحث عن هؤلاء المشاركين.

جودة البيانات. حتى شركات المسوح المتخصصة قد لا يكون لديها فهم جيد لتقييم الأثر وقد لا تكون مؤهلة وموثوق بها على النحو الذي نأمل. قد يقوم القائمون على المقابلة بتزيف المعلومات أو تسجيلها بشكل خاطئ. يجب ألا يكون إتباع أساليب ضعيفة لجمع البيانات مقبولاً. إذا تم اكتشاف بيانات مبتدعة أو منخفضة الجودة، فمن المهم إخطار شركة المسح بعدم قبول هذه البيانات وضرورة إجراء عملية جمع البيانات مرة أخرى لضمان اتساقها بأعلى

المعايير. من أجل الحد من هذه الحالات وكشفها، تأكد من وجود فريق مراجعة مستقل للإشراف على عملية جمع البيانات. في الغالب يتم مراجعة ١٠-١٥ بالمائة من المسوح لضمان وجود المستجيبين وجمع البيانات بشكل دقيق. عند العثور على مشكلات، قد يضطر إلى إعادة تدريب بعض القائمين على المسح أو فصلهم.

فقدان البيانات. قد يحدث ذلك في حالة فقد الاستبيانات المكتملة أو سرقة أجهزة الحاسوب أو تعطلها. لذا يتعين الحصول على نسخة أخرى من البيانات المحفوظة على أجهزة الحاسوب. في الميدان، يجب جمع المسوح من مجري المقابلة في أسرع وقت، مرتين أو ثلاثة مرات أسبوعياً، إذا أمكن، وذلك لحماية البيانات من الفقدان. إذا تم فقدان البيانات بالكامل، يفضل العودة مرة أخرى وإعادة جمع البيانات، أي زيارة الأفراد الذين تم إجراء المسح عليهم مرة أخرى مع التوضيح لهم أننا في حاجة إلى طرح الأسئلة عليهم مرة أخرى، الأمر الذي قد يكون مزعجاً للمستجيبين ومكلفاً للبرنامج.

إدخال البيانات. يجب إدخال البيانات فور جمع المسوح. الأمر الذي يسمح بتحديد المشكلات وتصحيحها في الميدان بشكل سريع. علاوة على ذلك، غالباً ما تقع أخطاء أثناء عملية جمع البيانات. تسمح معظم حزم الحاسوب الخاصة بإدخال البيانات (ولا تتطلب) بإدخال مزدوج حيث يتم إدخال كل قيمة مرتين. يتم الحد من الأخطاء المطبعية من خلال استخدام الهواتف المحمولة أو أجهزة المساعد الرقمي الشخصي PDAs أو أجهزة الكمبيوتر المحمولة أو أجهزة الكمبيوتر اللوحية لإدخال البيانات.

افتراضات خاطئة. قد تظهر الافتراضات الرئيسية لتصميم التقييم المختار عدم مواءمتها. فمن خلال إجراء اختبارات التحقق والتزيف (أنظر ملحق ٣)، يمكننا الكشف عن هذه الحالات من التحليل الأولي واتخاذ الإجراء المناسب، بما في ذلك تعديل استراتيجية التقييم. من أجل الحد من فرص إبطال تصميمنا، فمن الضروري أن يحافظ كل من العاملين بالتقييم والبرنامج على الاتصال والتعاون عن كثب لضمان تماشى كل من عملية التسجيل بالبرنامج وجمع البيانات مع متطلبات التقييم.

إجراء مسح وتحليل متابعة

التناقض. يُعد التناقض من المشكلات الكبرى التي تواجهها الدراسات ويمكنها خفض قيمة النتائج بشكل كبير. ويتضح إن الوقاية خير من العلاج. إن الحصول على بيانات الاتصال خلال المرحلة الأولية، وتقديم حوافز للشباب للمشاركة في المسح وكذا استخدام مسوح المتابعة يحد من درجة التناقض. إذا تعرض البرنامج إلى درجة عالية من التناقض، رغم الجهود المبذولة لمنع، يتمثل أحد أساليب تخفيف التناقض في اختيار عينة عشوائية من الأفراد الذين لم يتم تحديد مواقعهم والبحث عنهم بشكل مكثف. إذا تم العثور على هؤلاء الأفراد فقد يمثلون من لم يتم متابعتهم. وفي النهاية، حيث إنه لا يمكن تجنب حدوث بعض التناقض، يمكن أخذ هذا التناقض في الاعتبار عند تحديد عينة التقييم. إن إعداد عينة تزيد عما يجب أن تكون عليه بنسبة ١٠-٢٠٪ يسمح بالحصول على عدد أكبر من إجابات المسح للعثور على نتائج ذات دلالة إحصائية حتى في حالة حدوث تناقض (رغم أن هذا الأمر لا يلغي التحيز المحتمل من التناقض).

عدم الامتثال. بالإضافة إلى التناقض، في بعض الحالات لا يمثل الأشخاص لمعايير اختيار البرنامج بشكل كامل. على سبيل المثال، قد لا يحضر الشباب الذين تم اختيارهم للمشاركة إلى البرنامج التدريبي، بينما قد يحضر هؤلاء الذين تم اختيارهم في مجموعة المقارنة. ومن ثم فإجراء مقارنة دقيقة للنتائج بين مجموعة المعالجة ومجموعة المقارنة لن يمثل الأثر الفعلي للبرنامج. طالما أن هذه الحالات محدودة العدد ونعلم تحديداً أفراد مجموعتي المعالجة والمقارنة وماهية التدريبات التي تلقوها (من خلال سجلات البرنامج)، يمكن إجراء تصحيحات لعدم الامتثال وذلك باستخدام تقنيات إحصائية، مثل تقدير «المعالجة على أساس من تم معالجتهم»، والتي سيتمكن المقيم من حسابها.

تقييم الصندوق الأسود. هناك مشكلة أخرى شائعة كذلك في عملية المتابعة ألا وهي الافتقار إلى معرفة إلى أي مدى تم تنفيذ البرنامج بشكل جيد. الأمر الذي ينتج عنه تقييمات لا تعزي تغييرات ملحوظة (أو عدم حدوثها) لتصميم البرنامج أو تنفيذه. ومن الحلول الشائعة هي دمج النتائج التي تم الحصول عليها من نظام المتابعة وتكملة تقييم الأثر بتقييم سير العملية (انظر كذلك إلى الأساليب المختلطة في بند ٨).

نشر النتائج

الاستخدام المحدود لنتائج التقييم. إذا لم يتم مشاركة نتائج التقييم بشكل كاف مع أصحاب المصالح على المستوى الداخلي والخارجي، تتأثر أهداف التعلم الأساسية الخاصة بتقييم البرنامج بوجه خاص وقطاع سبل عيش الشباب بوجه عام. وإحدى طرق التغلب على هذه المشكلة يتمثل في تحديد استراتيجية النشر منذ بداية التقييم على أن يعمل أحد العاملين بالبرنامج عن كثب مع فريق عمل التقييم. ومن ثم، يكون هناك أحد الأشخاص الرئيسيين على دراية وفهم بالتقييم ويكون في وضع جيد يمكنه من تنفيذ بعض من نتائج التقرير.

النقاط الرئيسية

١. قد تكون عملية إجراء تقييم الأثر مهمة مكلفة وتستغرق وقتاً طويلاً وتواجهها الكثير من العثرات المحتملة. ومن ثم يكون من الضروري تجميع فريق على مستوى عالٍ من الجودة يمكنه العمل على التقييم خلال فترة زمنية ممتدة.
٢. خطة التقييم هي أكبر أول منتج لتقييم الأثر. وهي تظهر الإستراتيجية على كيفية تقييم التدخل، بما في ذلك منهجية البحث وحجم العينة وخطة جمع البيانات وغيرها من العناصر الأخرى.
٣. يفرض إجراء المقابلات مع الأطفال والشباب تحديات محددة بدءاً من الحصول على موافقة من الوالدين وحتى استخدام اللغة التي تناسبهم، ومن ثم يُنصح بتعيين شخص متخصص في هذه المسوح. علاوة على ذلك، يمكن أن تطرح التقييمات أسئلة أخلاقية، لذا يتعين التوجه إلى مجلس المراجعة المؤسسية للموافقة على تصميم التقييم والمسح.
٤. ينصح بشدة إجراء مسح أولي والذي من شأنه توفير معلومات قيمة تفيد المرحلة الخاصة بتصميم البرنامج وتسمح لنا بالتأكد من جدوى التصميم الذي تم اختياره للتقييم.
٥. يجب التفكير ملياً في توقيت جمع بيانات المتابعة من أجل رصد النتائج ذات الأهمية، حيث تظهر بعضها في الأجل القصير يحتاج بعضها الآخر إلى أعوام لتتجسد.
٦. يعد نشر نتائج التقييم، سواء كانت إيجابية أو سلبية بشكل واسع من الأمور الحاسمة. حيث توفر مشاركة النتائج مع أصحاب المصالح على المستوى الداخلي والخارجي والدولي أساساً للتعلم والتغذية الراجعة.

دراسة حالة NUSAF: تنفيذ تقييم الأثر

بدأ تقييم برنامج فرص الشباب الخاص بدراسة NUSAF في حزيران/يونيو ٢٠٠٧ وتم استكماله في أيار/مايو ٢٠١١ مع وضع التقرير الختامي. قام البرنامج بتوزيع الأموال على المشاركين في الفترة من آب/أغسطس إلى أيلول/سبتمبر ٢٠٠٨. واشتمل التقييم على إجراء مسح أولي في مستهل عام ٢٠٠٨ ومسح متابعة في أواخر عام ٢٠٠٩ فضلاً عن مسح ختامي في أواخر ٢٠١٠-مستهل ٢٠١١. وقد غطت جميع المسوح فئة المشاركين كافة.

المهمة	٢٠١١	٢٠١٠	٢٠٠٩	٢٠٠٨	٢٠٠٧
البرنامج					
تصميم البرنامج					
اختيار المقاطعة وتوزيع ميزانية المشروع					
تطوير متغيرات المعالجة					
التقييم المكتبي والميداني للمقترحات المؤهلة					
اختيار مجموعتي المعالجة والمقارنة					
المكون ١: نقل الأموال إلى مجموعات الشباب					
المكون ٢: بناء القدرة					
تقييم الأثر					
إعداد خطة وأسلوب التقييم					
تقدير ميزانية شركة المسح وبروتوكولات المسح الموصى بها					
وضع مسودة لتطوير المسح واختباره (باللغة الإنجليزية)					
اختيار مجموعتي المعالجة والمقارنة					
تعيين موظفي التعداد للمرحلة الأولى					
تدريب موظفي التعداد للمرحلة الأولى					
ترجمة المسح وإجراء اختبار تمهيدي له للمرحلة الأولى					
إجراء مسح أولي لمجموعة المعالجة والمقارنة					
إدخال البيانات					
تحليل البيانات					
تقرير عن بيانات المرحلة الأولى					
تعيين موظفي التعداد لمرحلة التتبع					
تدريب موظفي التعداد لمرحلة التتبع					
ترجمة المسح وإجراء اختبار تمهيدي له لمرحلة التتبع					
إجراء مسح تتبعي لمجموعة المعالجة والمقارنة					
إدخال البيانات					
تحليل البيانات					
التقرير ببيانات التتبع					
تعيين موظفي التعداد للمرحلة الختامية					
تدريب موظفي التعداد للمرحلة الختامية					
ترجمة المسح وإجراء اختبار تمهيدي له للمرحلة الختامية					
إجراء المسح الختامي لمجموعة المعالجة والمقارنة					
إدخال البيانات					
تحليل البيانات					
التقرير الختامي					
نشر البيانات					

المصدر: باتمان ومارتينيز وفيالا (٢٠٠٩).

قراءات رئيسية

Baker, J. 2000. *Evaluating the Impact of Development Projects on Poverty: A Handbook for Practitioners*. Washington, DC: The World Bank. (Chapter 2 is relevant to this note.) <http://siteresources.worldbank.org/INTISPMA/Resources/handbook.pdf>

Bamberger, M., Rugh, J., and Mabry, L. 2006. *Real World Evaluation: Working under Budget, Time, Data and Political Constraints*. Thousand Oaks: Sage Publications. (See chapters 3–8.) <http://realworldevaluation.org/>

Gertler, P., Martinez, S., Premand, P., Rawlings, L., and Vermeersch, C. 2011. *Impact Evaluation in Practice*. Washington, DC: The World Bank. (See chapters 10–13.) <http://www.worldbank.org/ieinpractice>

بند ٧: زيادة أهمية تقييم الأثر

بالرغم من أن تقييم الأثر الكمي المعياري في حد ذاته يمثل قيمة كبيرة لبرنامجنا ومؤسستنا، إلا إنه يوجد العديد من الخيارات المتاحة التي يمكنها أن ترفع من جودة التحليل وتزيد من أهمية النتائج. أولاً، يمكن لتقييمات الأثر أن تجيب عن مختلف أسئلة البحث، بما فيها متوسط تأثيرات البرنامج وأكثر من ذلك. ثانياً، لدى دمجها مع الأساليب النوعية، يمكن أن تصبح تقييمات الأثر أكثر إفادة وتغلب تلك التي تتوافر لدينا في حالة الاعتماد على الأساليب الكمية فحسب. وأخيراً، يمكن تقوية نتائج تقييم الأثر من خلال إجراء مزيد من التحليلات، فعلى سبيل المثال، يمكن قياس الفوائد الكلية للبرنامج مقابل التكاليف الكلية (من خلال تحليل المنافع والتكاليف). يعرض هذا البند الأخير نبذة مختصرة عن هذه الأدوات بحيث يمكن للممارسين أن يحققوا الاستفادة القصوى من تقييم الأثر الخاص بهم.

قياس مجموعة متنوعة من الآثار

يجيب تقييم الأثر الأساسي عن السؤال التالي؛ «هل عمل البرنامج بنجاح؟». بمعنى آخر، هل أثر على النتائج موضع الاهتمام على النحو المذكور والمحدد على البرنامج وأهداف التعلم التي وضعناها؟ السؤال عما إذا كان أثر البرنامج ككل مهما سؤال هام، ولكنه بالطبع ليس هو السؤال الوحيد الذي نود طرحه.

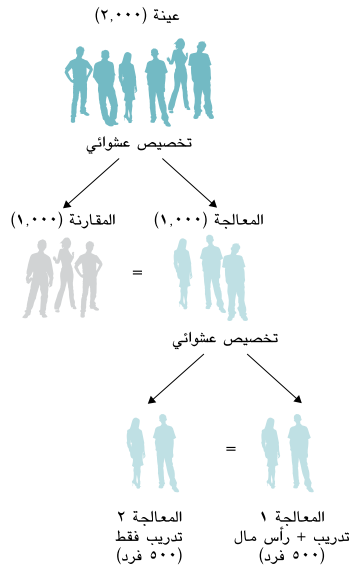
أولاً، من المفيد أن يكون لدينا صورة أكثر دقة وتفصيلاً حول الأثر الفعلي للبرنامج. وهذا يتضمن الحصول على إجابة للأسئلة التالية مع فهمها واستيعابها بصورة أفضل:

- هل تباينت النتائج على مستوى مختلف مجموعات المستفيدين (مثال، استفاد الأولاد من البرنامج، بينما لم تستفد الفتيات)؟
- ما هو الأثر قصير الأجل مقابل الأثر طويل الأجل لهذا التدخل؟
- هل لهذا البرنامج آثار جانبية سواء كانت إيجابية أو سلبية؟ هل كانت هناك أية نتائج سواء كانت مقصودة أو غير مقصودة خارج نطاق المجموعة المستهدفة فعلياً؟
- ثانياً، يمكن أن نكون كذلك مهتمين باختبار التصميمات الشاملة (CCDs) وكذلك اختبار كيف تتغير فاعلية برنامجنا كلما أجرينا تغييراً على التصميم. تتقصى التصميمات الشاملة الأسئلة التالية:
- هل تصميم أحد البرنامج أكثر فاعلية من غيره؟ قد نرغب في مقارنة التدخلات البديلة (على سبيل المثال، تقديم منح بداية المشروع مقابل قروض بداية المشروع لأصحاب المشروعات من الشباب)، أو اختبار أكثر الأنماط فاعلية للجمع بين مكونات البرنامج (التدريب فقط، التدريب والبرامج التدريبية، التدريب مع البرامج التدريبية والتوجيه).
- ما هي الجرعة الأنسب والأكثر فاعلية المناسبة لتقديم أنشطة البرنامج؟ على سبيل المثال، ما هو عدد الساعات الأنسب للتدريب؛ ٢٠ أم ٥٠ أم ١٠٠ ساعة؟
- إذا تم تصميم تقييمات الأثر بالشكل الصحيح، يمكنها أن تجيب عن هذه الأسئلة كافة، إلا إنه من الصعب الإجابة عن جميع الأسئلة باستخدام تقييم أثر واحد فقط. ولأن لكل تدخل أولويات وأهداف تعلم مختلفة، يمكن أن نصمم تقييم الأثر على النحو الذي يجيب عن الأسئلة الأكثر أهمية بالنسبة لبرنامجنا. ومن خلال تناول مجموعة أوسع نطاقاً من الأسئلة، يمكن أن نحسن أهمية نتائج التقييم. ولكن، من المهم أيضاً أن نفهم ونعي جيداً أنه يلزم توفير بيانات إضافية لتقييم الأسئلة التفصيلية والتصميمات الشاملة (أنظر جدول ٨-١).

السؤال	الوصف	متطلبات البيانات الإضافية	عينة من نتيجة وتفسير التقييم
ما هو الأثر الكلي للبرنامج على أ و ب وج في المجموعة X؟ في السياق Y؟	هذا هو سؤال تقييم الأثر المعياري.	لا يوجد (جمع البيانات المعيارية يكون بناء على الأسلوب الذي يتم اختياره).	متوسط أثر البرنامج التدريبي على دخل الشباب هو زيادة ٢٠ دولاراً شهرياً. البرنامج ذو أثر إيجابي على الدخل.
هل تتباين النتائج على مستوى مجموعات مجتمع الدراسة؟	يكون للتدخلات غالباً تأثيرات متباينة على مستوى المجموعات (عدم التجانس). وقياس متوسط الأثر تختفي معالم هذه الاختلافات؛ لذا فعلينا أن نفصل الآثار الخاصة بكل مجموعة.	- المعلومات الاجتماعية الديموغرافية الخاصة بالمشاركين ومجموعة المقارنة (العمر والنوع ومستوى الدخل وغيرها). - كي تتمكن من تنفيذ النتائج، يتعين زيادة عدد الأفراد الذي يشملهم التقييم (حجم العينة) مع كل فئة معلومات يتم تحليلها.	متوسط الزيادة في الدخل ٤٠ دولاراً بالنسبة للأولاد و٠ بالنسبة للفتيات. يستفيد الشباب الأكبر سناً أكثر من الشباب الأصغر سناً (بمتوسط ٣٠ دولاراً مقابل ١٠ دولاراً). لذلك، فإن فاعلية البرنامج لا تقف على قدم المساواة بالنسبة لكافة المشاركين. وعلينا أن نفهم لماذا يختلف مستوى استفادة المجموعات عن بعضها البعض ويمكن طريقة الاستهداف في البرنامج وتصميمه بالشكل الذي يتلاءم مع مجموعات بعينها.
ما هو الأثر قصير الأجل مقابل الأثر طويل الأجل للبرنامج؟	قد لا يكون التغيير في النتائج ثابتاً على مدار الوقت. فقد تتلاشى الآثار قصيرة الأجل بينما قد لا تظهر الآثار طويلة الأجل إلا بعد عدة سنوات من نهاية التدخل.	البيانات على مدار فترة زمنية ممتدة (عملياً غالباً ما يقصد بذلك متابعة مجموعات المعالجة والمقارنة لسنوات عديدة).	في نهاية البرنامج، لاحظنا أن متوسط دخل المشاركين الشهري - ٥ دولارات (خسارة) مقارنة بمجموعة المقارنة. وبعد عامين من البرنامج، كان متوسط الزيادة في الدخل الشهري لأفراد مجموعة المعالجة ٢٠ دولاراً. لم يكن هؤلاء الذين اشتركوا في التدريب قادرين على العمل بنفس قدرة أقرانهم خلال فترة التدريب، ولذلك انخفض دخلهم. إلا إنه مع مرور الوقت، حقق التدريب مسعاه واستطاع المشاركون تأمين دخل أعلى من أقرانهم الذين لم يشاركوا. وبهذا فإن النظر إلى النتائج قصيرة الأجل فحسب قد يؤدي إلى نتائج مضللة.
هل البرنامج له تأثيرات جانبية؟	قد يكون للبرنامج تأثيرات غير مباشرة على غير المشاركين (إيجابية أو سلبية).	بيانات بخلاف مجموعة المعالجة والمقارنة، تتضمن أفراد العائلة والمجتمع المحيط بالمجموعة.	لم يتمتع المشاركون فحسب بمتوسط الزيادة في الدخل الشهري التي قدرت بـ ٢٠ دولاراً، بل أن جيرانهم أيضاً تمتعوا بزيادة بنحو ٥ دولارات زيادة في دخلهم الشهري؛ فيما يبدو أن المشاركين قد نقلوا خبراتهم وشاركوا مع الآخرين.
هل تصميم البرنامج أ أو تصميم البرنامج ب أكثر فاعلية؟	حتى أفضل تصميمات البرامج يشوبها غالباً الغموض. فيمكن أن تجد أسئلة تتعلق بمقارنة التدخلات البديلة أو الجمع بين مكونات البرنامج.	- العديد من مجموعات المعالجة (إحداها) تلقت التصميم أ، وأخرى تلقت تصميم ب، وهكذا). - يتعين زيادة عدد الأفراد الذي يشملهم التقييم بصورة كافية كي نتأكد من تكوين أكثر من مجموعة معالجة واحدة وكذلك مجموعة مقارنة.	متوسط الزيادة في الدخل ٥ دولارات بالنسبة لهؤلاء الذين تلقوا التدريب، و٣٠ دولاراً بالنسبة لهؤلاء الذين تلقوا التدريب والبرامج التدريبية. لذلك فإن تقديم الخبرة العملية بالإضافة إلى التدريب تزيد من الأثر بشكل بالغ.
ما هي الجرعة الأنسب والأكثر فاعلية لهذا التدخل؟	ليس المزيد أفضل دائماً؛ فإيجاد التوازن الصحيح لحجم الخدمات التي يتم تقديمها أمر هام للغاية لتحقيق الأثر بأعلى استفادة ممكنة من ناحية ولتحد من التكاليف من ناحية أخرى.	- العديد من مجموعات المعالجة (إحداها) تلقت التصميم أ، وأخرى تلقت تصميم ب، وهكذا). - يتعين زيادة عدد الأفراد الذي يشملهم التقييم بصورة كافية كي نتأكد من تكوين أكثر من مجموعة معالجة واحدة وكذلك مجموعة مقارنة.	متوسط الزيادة في الدخل ٠ دولاراً بالنسبة لهؤلاء الذين تلقوا تدريباً لمدة شهر واحد، و٢٠ دولاراً بالنسبة لهؤلاء الذين تلقوا تدريباً لمدة ثلاثة أشهر و٢٠ دولاراً بالنسبة للذين تلقوا تدريباً لمدة ستة أشهر. على الرغم من أن التدريب لمدة شهر واحد لم يكن كافياً، إلا أن التدريب لمدة ستة أشهر لم يكن له فوائد إضافية مقارنة بالتدريب لمدة ثلاثة أشهر. ومن الواضح أن الفترة المثلى لتلقي التدريب هي ما يقرب من الثلاثة أشهر.

تساعد التصميمات الشاملة في التعرف على ما يزيد عن مجرد معرفة الأثر الكلي للبرنامج؛ فإنها تقيم كذلك أداء الخصائص الخاصة بالبرنامج وتحدد لماذا تعمل هذه الخصائص أو لا تعمل بشكل صحيح. على سبيل المثال، قد يشمل أحد البرامج تدريباً على المهارات المهنية ومهارات تنظيم المشاريع، مثل التجارة والخيطة، فضلاً عن رأس مال صغير لبداية المشروع. قد يكون تقديم المنح المالية أمر باهظ التكلفة أو صعب لأسباب سياسية، وهنا قد يتساءل مدير المشروع حول ضرورة رأس مال بداية المشروع وما إذا كان يتوجب تقديمه أو أنه يمكن المشاركين تنفيذ التدريب دون الحصول على رأس المال. يمكن للتصميم الشامل مساعدتك في تحديد أفضل تصميم للبرنامج في هذه الحالة. تتطلب التصميمات الشاملة على الأقل مجموعتي معالجة تتلقى تركيبات وجرعات البرنامج المختلفة. يمكن مقارنة المجموعتين بعد ذلك عند الاستبيان الختامي، ويكون الفرق بين المجموعتين هو الأثر الخاص بتصميم معين عن غيره. إذا طبقنا المثال أعلاه، قد يجري البرنامج تقييماً على عينة تحتوي على ٢,٠٠٠ مشارك يتم تخصيصهم بشكل عشوائي للانضمام في مجموعتي المعالجة والمقارنة. يمكن بعد ذلك إعادة تخصيص مجموعة المعالجة بشكل عشوائي لتكوين مجموعتي معالجة. في المعالجة الأولى، يتم تقديم التدريب ورأس مال بداية المشروع. وفي المعالجة الثانية، يتم تقديم التدريب فحسب (أنظر شكل ٨-١).

شكل ٨-١ مخطط لتقييم الأثر بمكون تصميمي شامل



يمكن حينها تحديد أثر توفير رأس مال من خلال مقارنة أفراد المعالجة الأولى بالمعالجة الثانية في التقييم الختامي. ويتم تحديد أثر التدريب من خلال مقارنة أفراد مجموعة المعالجة الثانية بأفراد مجموعة المقارنة.

استخدام مناهج الأساليب المختلطة

من الهام أن نأخذ في اعتبارنا حدود أساليب تقييم الأثر الكمي. إذا تم الاعتماد عليها بمفردها، سنقع في فخ خطير يصعب معه فهم نتائج البرنامج المعقدة أو تفسير الآثار التي قد تظهر تفسيراً صحيحاً. ولكي نحصل على فهم راسخ صحيح لديناميكيات التدخل ولكي نتمكن أيضاً من شرح لماذا تعمل الأشياء على نحو معين، فمن الضروري دمج تقنيات تقييم الأثر في إطار من متابعة قوي وتقييم سير العملية. وبشكل عام، نحن نؤمن بأن استخدام الأساليب المختلطة التي تضم بشكل واضح كلا من الأساليب النوعية والكمية في تصميم تقييم الأثر يمكن أن تطور بشكل بالغ من عملية التعلم في برامجنا والتعرف عليها.

وكما أشار بامبيرجير وراو وولكوك (٢٠١٠، ص ٦-٧)، وكذلك ليو وفيسين (٢٠٠٩)،

يوجد العديد من الطرق تعمل من خلالها الأساليب على تقوية تقييم الأثر الكمي:

- عادة لا تجمع تقييمات الأثر الكمية معلومات عن جودة تنفيذ البرنامج، في حين يعد فهم عملية التنفيذ أمراً بالغ الأهمية لفهم كيف أثر تنفيذ البرنامج على نتائجه وكذلك كي يتم تفسير النتائج بشكل صحيح والتعرف على ما إذا كانت النتائج غير المرصية ناشئة عن نقاط ضعف في تصميم البرنامج أم في تنفيذه. إذن، فالمتابعة الراسخة شرط أساسي للتقييم الفاعل ويمكن إتمامه باستخدام أدوات إضافية لتحليل سير العملية مثل مقابلات المبلغين الأساسيين والملاحظة المباشرة للمشاركين ومجموعات النقاش البؤرية/المركزة.
- يمكن كذلك أن يساعد إدراج الأساليب النوعية في فهم كيف ولماذا يتباين ويختلف تأثير التدخل على مستوى الأفراد المستهدفين. وعلى الرغم من أن الأساليب النوعية يمكن تصميمها للتعرف على عدم تجانس الأثر على مستوى المجموعات، إلا إنها لا يمكن أن تقدم تفسيراً واضحاً حول سبب حدوث عدم التجانس هذا.
- وعلى الرغم من أن التصميمات الكمية وحدها قد لا تستطيع أن تكشف نطاق الظروف الداخلية التي تم فيها تنفيذ كل مشروع، يمكن للأساليب المختلطة أن تساعد في تقديم تحليل مفصل حول السياق وتوثيق الاختلافات في جودة التنفيذ أو سرعته على مستوى مواقع البرنامج بأسره. ويمكن لهذه المعلومات النوعية بدورها شرح الاختلافات المحتملة في نتائج البرامج في المناطق الجغرافية المختلفة.
- تكون العديد من نتائج تدخلات سبل عيش الشباب (مثل الصحة النفسية أو التمكين أو العلاقات المنزلية) غاية في التعقيد ومتعددة الأبعاد وقد لا يمكن التعرف عليها من خلال الأساليب الكمية. تسمح الأساليب المختلطة بتتبع المؤشرات النوعية ومن ثم تقدم تحليلاً حول دراسة الحالة المختارة للمساعدة في فهم ديناميكيات ونتائج التدخل بصورة أفضل. على سبيل المثال، المقابلات النوعية الهيكلية أو شبه الهيكلية حيث يجد المشاركون الحرية في مشاركة قصصهم الواقعية التي تخرج عن نطاق فئة المعلومات الكمية. يمكن أن تساعد في الإلمام بفهم جيد حول أثر البرنامج. قد تكون الأساليب النوعية كذلك مناسبة في حالة جمع المعلومات حول الموضوعات الحساسة مثل تلك المتعلقة بالصحة الإنجابية أو بالعنف.
- ويمكن أن تساعد الأساليب النوعية أيضاً في تحديد المؤشرات المناسبة في المقام الأول. على سبيل المثال، قد تأتي مجموعة النقاش البؤرية/المركزة بمعلومات هامة حول مخاوف المستفيدين وكيف يعتقدون أن التدخل قد يؤثر على حياتهم.
- وبشكل عملي، تأخذ عملية دمج العناصر النوعية في تقييم الأثر أشكالاً عدة، منها أسئلة الاستبيان ذات النهايات المفتوحة ودراسات الحالة والمقابلات المعمقة المختارة ومجموعات النقاش البؤرية/المركزة والأدوات التشاركية مثل أسلوب «أهم تغيير على الإطلاق» «Most Significant Change» (أنظر دافيز ودارت ٢٠٠٥)، وملاحظة المشاركين وما شابه من أساليب. لمعرفة المزيد عن المتابعة والتقييم التشاركية، راجع كاتلي وزملائه (٢٠١٠)، وسابو فلورز (٢٠٠٨)، وباورز وتيفاني (٢٠٠٦) وغاولر (٢٠٠٥).
- وفي الوقت نفسه، لا يكون تطبيق المعلومات النوعية وحده كافياً للتعرف على آثار البرنامج. حيث يسمح تطبيق الأساليب المختلطة لنا بأن نجتمع بين نقاط القوة الخاصة بكل من أدوات التقييم الكمي والنوعي مع تجنب نقاط ضعفها واستبعادها، مما يسمح بتوفير تصميم تقييمي أقوى بشكل عام (يقدم مربع ٨-١ مثلاً على ذلك). في الواقع، إن الجمع بين تطبيق العديد من الأساليب البحثية كذلك يزيد من موثوقية نتائجنا وصدقها.
- بالرغم من ذلك، من المهم جداً ملاحظة أن تطبيق تصميمات الأساليب المختلطة يمكن أن تتطلب تكاليف إضافية ووقت إضافي فضلاً عن التحديات اللوجستية. بالإضافة إلى ذلك، عادة ما تستطيع الانقسامات المهنية بين التخصصات والباحثين أن تجعل «بناء فريق متعدد التخصصات قضية تستنزف الوقت وتشكل تحدي» (بامبيرجير وراو وولكوك ٢٠١٠ - ص ١٧).

في تقييم منظمة إنجاز الشباب لمنهج أمتنا، قام المقيمون بالجمع بين تشكيلة من أساليب البحث الكمية والنوعية. ويعد أمتنا أحد البرامج المتعددة لمنظمة إنجاز الشباب العالمية والذي تم توزيعه في كافة أنحاء العالم على المدارس الابتدائية ويتضمن سلسلة من الدروس للطلاب الذين تتراوح أعمارهم من ٩ إلى ١١ عاماً والتي تبحث مسائل متعلقة بتنظيم المشروعات، والموارد اللازمة للمشروعات التجارية، والعولمة. ومن ناحية أخرى، اعتمد التقييم على تصميم تجريبي مع التخصيص العشوائي للطلاب بين مجموعتي المعالجة والمقارنة. فجاء طلاب مجموعة المقارنة من نفس الولايات والمناطق التي أتى منها طلاب مجموعة المعالجة، ولكن فصولهم الدراسية كانت موزعة بعشوائية لتلقي البرنامج بعد إتمام التقييم. وبالإضافة إلى ذلك، أجرى المقيمون عدة دراسات حالة، باستخدام مقابلات مع المعلمين والمتطوعين والعاملين بمنظمة إنجاز الشباب ومجموعات النقاش البؤرية/المركزة مع الطلاب ومراقبة الفصول الدراسية.

أظهرت نتائج التقييم الكمي بعض الآثار الإيجابية على معرفة الطلاب بالمحتوى المتصل بتنظيم المشروعات والعولمة، ولكن دون تأثير على مستويات المشاركة في المدرسة واكتساب مهارات القرن الحادي والعشرين. وعلاوة على ذلك، سمحت الأدوات النوعية بفهم عميق للآليات المطبقة في جوهر عملها. من ناحية أخرى، أكدت الأدوات النوعية على جودة التنفيذ، في ظل إدارة غالبية المحاضرات طبقاً لإرشادات منظمة إنجاز الشباب وارتفاع مستويات الرضا لدى الطلاب والمعلمين والمتطوعين. ومع ذلك، فقد أشاروا أيضاً إلى التحديات التي تواجه تنفيذ برنامج جيد، ومنها على سبيل المثال، عدم كفاية الوقت للمتطوعين لتغطية كافة المحتويات. وأخيراً، اقترحت نتائج التقييم النوعي طرقاً لتحسين البرنامج، منها إطالة وقت المحاضرات وتخفيف صعوبة المفردات المستخدمة وتقديم مزيد من التوجيه في التدريس للمتطوعين.

المصدر: بحوث آرام سي (٢٠٠٩).

تحليل المنافع والتكاليف وتحليل فاعلية التكاليف

في حالات كثيرة، قد تستخدم المنظمات إستراتيجيات مختلفة لمواجهة نفس المشكلة. فعلى سبيل المثال، لزيادة فرص العمل، قد يتطلب الأمر تحسين تقديم الإرشاد الوظيفي أو تحسين التدريب. وحتى عندما يتم إتباع إستراتيجية واحدة، يمكننا اتخاذ مناهج مختلفة لتنفيذها، مثل استخدام مقدمي تدريب من القطاع العام أو من القطاع الخاص. وإذا ثبت أن هذه الإستراتيجيات العديدة سواء المتصلة بالبرنامج أو بالتنفيذ كان لها نفس الأثر، على سبيل المثال، إذا ثبت أن كل منها يحسن احتمال إيجاد وظيفة بعد التدخل بثلاثة شهور بنسبة ٥٠ بالمائة، فهل سنكون سعداء بنفس القدر لتنفيذ منهج دون الآخر؟ لا على الأرجح. فلا يكفي أن تعرف صلاحية تدخل ما، أو لمن يصلح، أو في أي نطاق يصلح، لأننا نحتاج أيضاً لمعرفة تكلفة هذا التدخل.

ومعرفة تقدير واقعي للتكاليف، يسمح لنا هذا بدوره بالإجابة عن الأسئلة التالية:

- كيف يمكننا الاختيار بين البدائل؟ ما البرنامج الأكثر فاعلية للتكاليف بالنظر إلى مستوى معين للأثر؟
- هل يمكننا التوسع؟ إذا كانت التكاليف مرتفعة، فإنه من غير المرجح أننا سوف تتمكن من التوصل لعدد كبير من المستفيدين.
- هل أي تدخل دائماً أفضل من لا شيء على الإطلاق؟ إذا كان إجمالي التكاليف يفوق إجمالي منافع وفوائد البرنامج، إذن فمن الأفضل استخدام الموارد في مجال آخر.

أدوات تحليلية

للإجابة عن الأسئلة أعلاه، هناك أداتان مستخدمتان بشكل شائع لذلك وهما تحليل فعالية التكاليف (CEA) وتحليل المنافع والتكاليف (CBA):

تحليل فعالية التكاليف يحدد تكلفة البرنامج بأكمله ويربط هذه التكاليف بتدابير محددة للمخرجات أو النتائج (٥٠٠ دولار أمريكي لكل شخص مدرب، ولكل وظيفة تم إيجادها، ولكل حالة داء نقص المناعة المكتسبة (الإيدز) تم منعها، وما إلى ذلك). وبناء عليه، يطلعنا تحليل فعالية التكاليف على مقدار المخرجات أو النتائج التي حصلنا عليها مقابل كل دولار تم صرفه، وبذلك نتعرف على التخصيص الأكثر كفاءة للموارد عندما نقارن البرامج البديلة بنفس المعيار (أنظر جدول ٨-٢).

جدول ٨-٢ تقديرات فعالية التكاليف لبرامج حركة مناصرة الشباب Jóvenes

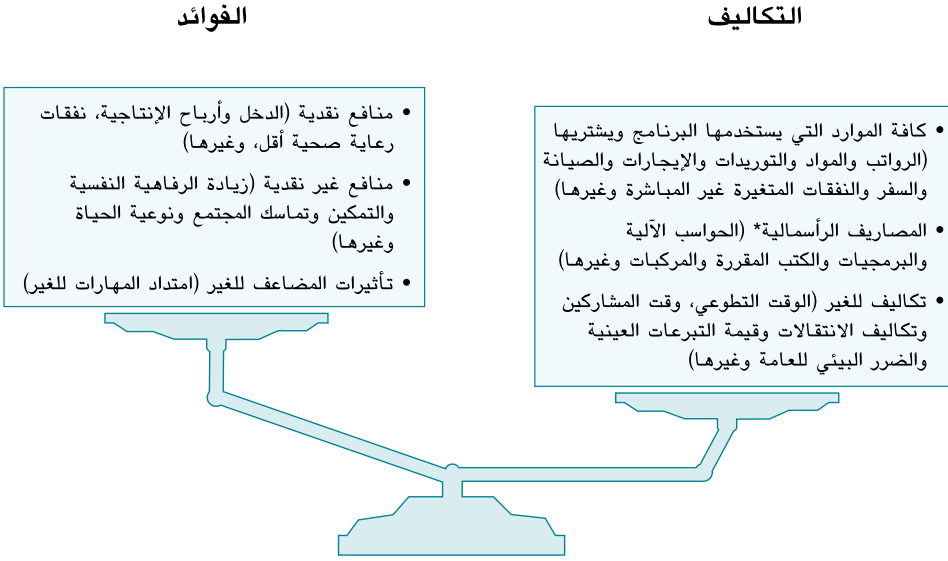
البلد	البرنامج	التكلفة لكل مشارك (في ٢٠٠٥ بالدولار الأمريكي)
الأرجنتين	برويكتو خوفين Proyecto Joven	١١٥٩ دولار
تشيلي	تشيلي خوفين Chile Joven	٨٢٥ دولار - ١٠٥١ دولار
بيرو	بروخوفين PROJoven	٦٩٧ دولار

المصدر: بيتشرمان وآخرين (٢٠٠٧).

تحليل المنافع والتكاليف يحدد أيضا تكلفة البرنامج بالكامل ويقارن هذه التكاليف مقابل قيمة جميع منافع البرنامج بالدولار. وبمعرفة صافي منافع وصافي تكاليف التدخل، يمكن عندئذ حساب نسبة المنافع إلى التكاليف وتحديد العائد للمجتمع على استثمار المنظمة. فعلى سبيل المثال، تكون نسبة المنافع إلى التكاليف ٢: ١ إذا كان صافي الفوائد ١,٠٠٠ دولار لكل شخص وكان صافي التكاليف ٥٠٠ دولار. بصفة عامة، فإن تحليل المنافع والتكاليف يسعى لتحديد مدى كون المنافع تفوق التكاليف، أي إذا ما أصبح المجتمع أكثر غنى أو فقراً بعد القيام بهذا الاستثمار. يمكن استخدام كل من تحليل فعالية التكاليف وتحليل المنافع والتكاليف قبل التدخل أو أثناء البرنامج أو بعده. ومع ذلك، يوفر التحليل بأثر رجعي فقط للممارسين معلومات كاملة عن التكاليف والمنافع الفعلية لتحديد النجاح الكلي للتدخل. وفي الواقع، يعتبر تقييم الأثر شرطاً ضرورياً للحصول على تقدير ممكن الاعتماد عليه لمنافع البرنامج المباشرة وغير المباشرة.

حصر كافة المنافع والتكاليف

يتطلب تحليل فعالية التكاليف وتحليل المنافع والتكاليف حصر كافة التكاليف المعروفة للبرنامج وقياس كميتها ومقارنتها (وبالنسبة لتحليل المنافع والتكاليف، المنافع المعروفة) لكل شخص تأثر بشكل مباشر أو غير مباشر بالتدخل: المنظمة المنفذة للبرنامج، والمستفيدين من البرنامج، والحكومة، والغير (أنظر شكل ٨-٢).



* شراء المواد التي يزيد استخدامها عن عام واحد.

قد تكون محاولة تقييم كثير من المنافع غير الملموسة بالدولار أمراً صعباً وغير موضوعي، ويمكن أن تمثل تحدياً كبيراً، ولا سيما بالنسبة لتحليل المنافع والتكاليف وبناء عليه، يعتبر تحليل المنافع والتكاليف عادة أكثر نفعاً عندما يكون هناك أنواع متعددة للمنافع وإجماع حول كيفية تحديد قيمتها نقداً (معمل عبد اللطيف جميل للتطبيقات العملية لمكافحة الفقر ٢٠١١).

حساب صافي المنافع للمشاركين

يمكن قياس منافع البرنامج بآثارها على المشاركين الأفراد. فعلى سبيل المثال، إذا وجد أن برنامج تدريب مهارات قد زاد من الدخل بمعدل ١٠٠ دولار للشخص سنوياً في المتوسط. في مناطق كثيرة في أفريقيا جنوب الصحراء الكبرى، يمثل هذا مبلغاً ضخماً من المال وقد يمثل نجاحاً كبيراً للبرنامج.

إذا افترضنا أن البرنامج يكلف مليون دولار لتنفيذه، بمبلغ ٤٠٠,٠٠٠ دولار لإجراء التدريب و٦٠٠,٠٠٠ دولار كتكاليف متغيرة غير مباشرة متضمنة كافة رواتب الموظفين. وإذا وصل البرنامج إلى ١,٠٠٠ شخص، يتكلف البرنامج إذن ١,٠٠٠ دولار للشخص لتنفيذ البرنامج، بمبلغ ٤٠٠ دولار لإجراء التدريب و٦٠٠ دولار كتكاليف متغيرة غير مباشرة. فهل يجدر تنفيذ البرنامج؟ تعتمد الإجابة عن هذا السؤال على ثلاثة معايير.

أولاً، يجب أن يكون أثر البرنامج مساوياً أو زائداً عن أثر منح الأفراد نقد مساوي لتكلفة تنفيذ البرنامج. في المثال أعلاه، يجب مقارنة الأثر بنتائج إعطاء كل شخص ١,٠٠٠ دولار نقداً. هناك سيناريوهان محتملان لذلك. أولاً، قد يستخدم هذا الشخص المبلغ الـ ١,٠٠٠ دولار ويشتري التدريب بهذا المبلغ. وفي هذه الحالة، تكون تكلفة التدريب فقط ٤٠٠ دولار للشخص، والتي تنتج نفس العائد البالغ ١٠٠ دولار للشخص سنوياً. وبالتالي، يكون لدى الفرد ٦٠٠ دولار إضافية لاستخدامها كيفما يحب، وبهذا يكون أفضل حالاً من اشتراكه في البرنامج. وفي السيناريو الثاني، قد يستخدم الشخص الأموال في أمر آخر بخلاف التدريب، والذي يكون أقل نفعاً له على المدى الطويل، مثل إنفاق المال على السجائر أو السلع الاستهلاكية غير الضرورية الأخرى. وفي الحالة الأخيرة، يكون من الأفضل تنفيذ البرنامج.

العائد الاجتماعي على الاستثمار (SROI) هو تعديل لتحليل المنافع والتكاليف الذي يقارن المنافع المالية الإضافية بالنسبة للموارد المستثمرة. ويخصص قيم مالية تقديرية لجميع هذه النتائج التي حددها أصحاب المصالح والتي ليس لها قيم سوقية بصفة عامة. لمزيد من المعلومات، يرجى الرجوع إلى الموقع الإلكتروني لشبكة العائد الاجتماعي على الاستثمار: <http://www.thesroinetwork.org>

ثانياً، يجب أن يكون للبرنامج عائد مساوي أو أكبر من عائد تنفيذ البرامج الأخرى الجارية. هل من الممكن أن يحقق برنامج آخر نفس الأثر أو أثراً أكبر للشخص بتكلفة أقل؟ يتطلب هذا السؤال مقارنة النتائج بين خيارات البرامج المختلفة. ويعتبر البرنامج الذي لديه أكبر أثر ولكن تكلفته أقل هو أفضل برنامج كي يستمر تنفيذه.

أخيراً، يجب أن يكون صافي القيمة الحالية للعائد أكثر من تكلفة البرنامج. فالقيمة الحالية هي طريقة للتفكير في قيمة الأموال اليوم مقارنة بقيمتها في المستقبل. وباستخدام المثال الحالي، فأنا نأخذ قيمة الأموال التي يتم الحصول عليها سنوياً (أي مبلغ ١٠٠ دولار للشخص والذي يمثل العائد على التدريب) وتعديل قيمتها عبر فترة زمنية طبقاً لمعدل الخصم، والذي يساوي في معظم الحالات، معدل الفائدة المحلي. وصافي القيمة الحالية وهو ببساطة مجموع القيمة الحالية المعدلة عبر فترة زمنية. ويتمثل هذا في جدول ٨-٣، باستخدام معدل فائدة قدره ٢٠ بالمائة، وهو المعدل السائد في العديد من البلدان النامية.

جدول ٨-٣ القيمة الحالية وصافي القيمة الحالية لعائد سنوي قدره ١٠٠ دولار أمريكي

العام	القيمة الحالية
٠	١٠٠ دولار
١	٨٣ دولار
٢	٦٩ دولار
٣	٥٨ دولار
٤	٤٨ دولار
٥	٤٠ دولار
٦	٣٣ دولار
٧	٢٨ دولار
٨	٢٣ دولار
٩	١٩ دولار
١٠	١٦ دولار
صافي القيمة الحالية	٥١٧ دولار

في هذا المثال، فإن مبلغ ١٠٠ دولار بقيمة اليوم يتم تقييمه بمبلغ ١٠٠ دولار. ومع ذلك، بمعدل الفائدة الحالي الخاص بنا، فإن دخل اليوم البالغ ١٠٠ دولار سوف يساوي ١٦ دولار فقط في عشرة أعوام. فعلى مدى عشرة أعوام، ستصبح صافي القيمة الحالية لمبلغ ١٠٠ دولار مساوية لمبلغ ٥١٧ دولار. وعلى مدى عمر المشارك، سوف تكون صافي القيمة الحالية ٦٠٠ دولار على الأكثر. وبناء عليه، يتطور عائد الـ ١٠٠ دولار للشخص سنوياً لأقصى عائد بمبلغ ٦٠٠ دولار للشخص على مدى حياته.

طبقاً لهذه المعايير، وبقيمة الدولار بتاريخ اليوم، فإن مصاريف البرنامج (١٠٠٠ دولار) أكبر من المنافع التي تعود على الأفراد، على الرغم من أن تكلفة التدريب (٤٠٠ دولار) أقل من صافي القيمة الحالية للتدريب (٥١٧ دولار). وبناء عليه، ما لم يمكن حث الأفراد على التقدم لإكمال التدريب بأنفسهم دون الحاجة إلى ميزانية التكاليف المتغيرة غير المباشرة، أو ما لم تخفض ميزانية التكاليف تلك بشكل كبير، فإن قيمة التدريب ليست كافية لتبرر تنفيذ البرنامج.

حساب صافي المنافع للمجتمع

لغرض تقييم صافي المنافع للمجتمع، يلزم دراسة امتداد التأثير. ويشير امتداد التأثير إلى الآثار الإيجابية أو السلبية التي تكون للبرنامج على الأشخاص غير المشتركين بشكل مباشر في البرنامج. وهناك نوعان من الامتدادات وهما موضع اهتمامنا هنا: تأثيرات المضاعف، وتأثيرات الجوائز والكمية.

تأثيرات مضاعف

تحدث تأثيرات المضاعف عندما يفصح المشاركون في البرنامج عن مهاراتهم للآخرين الذين لا يرتبطوا رسمياً بالبرنامج. على سبيل المثال، قد يقوم رجل تم تدريبه في مهنة النجارة بتدريب زوج ابنته. فتقييم الأثر قد يقيس الأثر الذي يكون للبرنامج على المشارك فقط، وقد يفوته النتيجة التي كانت للبرنامج على زوج الابنة.

باستخدام المثال أعلاه، كان الأثر على النجار ١٠٠ دولار في السنة، بينما الأثر على زوج الابنة كان أقل بسبب انخفاض جودة التدريب، ولكن من الواضح أن تدريب شخص واحد قد حسن من آفاق سبل العيش لشخصين. وبناء عليه، فإن التكلفة للشخص أقل من مبلغ ١,٠٠٠ دولار الذي تم حسابه أساساً.

قد تكون المنافع غير المباشرة هامة وقد تبرر تكاليف البرنامج في بعض الحالات. ولغرض حصر هذه الامتدادات، يجب إعداد الخطط أثناء جمع البيانات في المرحلة الختامية للسؤال عن الآخرين الذين تلقوا التدريب أو الذين استفادوا من البرنامج.

تأثيرات الجوائز والكمية

على الرغم أن التدخلات قد تستهدف شرائح معينة من السكان والسوق المحلي، إلا أنها قد يكون لها تأثيرات على الاقتصاد الأكبر، والتي يشار إليها في أحيان كثيرة على أنها تأثيرات «التوازن العام». فعلى سبيل المثال، إذا، نتيجة لتدريباتنا المقدم في النجارة، كان هناك نجارين مهرة آخرين في الاقتصاد المحلي، فقد تقلل المنافسة بينهم الأسعار للمستهلكين. كما يمكن أن يكون للبرنامج امتدادات سلبية على الاقتصاد. على سبيل المثال، تدريب خياطين إضافيين في منطقة بها الكثير من الخياطين قد يؤدي إلى تدني الأسعار للغاية مما يخرج بعض الخياطين من السوق. وقد يخفف هذا التأثير بشكل كبير من أثر البرنامج.

هناك أيضاً تأثير غير مرغوب آخر قد ينتج عنه عواقب سلبية لغير المشاركين. فإذا حصل المشاركون في تدخل معين على ميزة تنافسية في سوق العمل، مثلاً، قد ينتج عن ذلك عدم عثور بعض الشباب على وظيفة على الرغم من أنه كان من الممكن أن يجدوا وظائف في غياب البرنامج. ويشير إلى هذه التأثيرات بشكل عام على أنها تأثيرات إزاحة.

بالنسبة لغالبية البرامج، تكون تأثيرات الجوائز والكمية على الأرجح ضئيلة للغاية وغير جذيرة بجمع البيانات عنها. ولكن قد ترغب البرامج الكبيرة في استكشاف طرق لحصر هذه التأثيرات بالاستعانة بخبير تقييم. وأحد الطرق الممكنة لحصر هذه التأثيرات تتمثل في عشوائية التدخل على مستوى المجتمع أو مستوى المنطقة.

النقاط الرئيسية

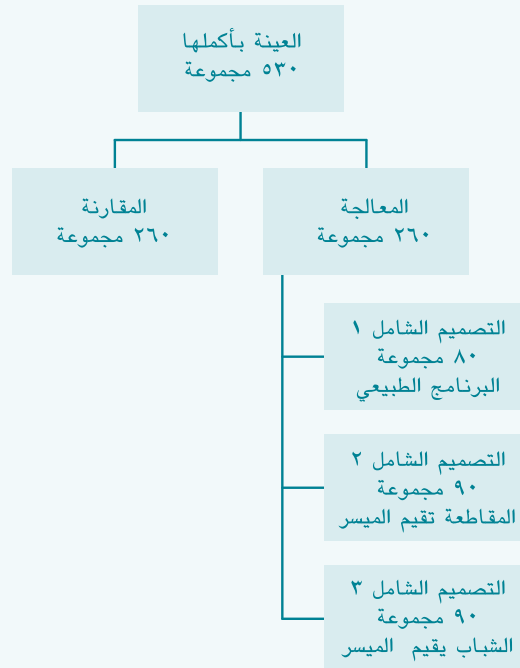
١. بناءً على أهداف التعلم للبرنامج وللمنظمة، يكون من الجدير استكشاف ما إذا يمكن تصميم تقييم الأثر لقياس أكثر من مجرد متوسط أثر البرنامج. ويمكن أن تتعلق أسئلة الأثر الإضافية هذه بعدم التجانس أو الأفق الزمني أو امتدادات الأثر أو الفاعلية النسبية لخيارات التصميم المختلفة.
٢. ومن المحبذ بشدة إدخال عناصر البحث النوعي في تقييم الأثر. فاستخدام الأساليب المختلطة يمنحنا فهماً أكثر شمولية ودقة لأثر البرنامج أو لغيابه.
٣. قد تكون المعلومات حول أثر البرنامج محدودة المنفعة ما لم نكن نعلم أيضاً بتكاليف تصميم وتنفيذ التدخل. وسوف يعتمد أي توسع في نطاق البرنامج على هذه المعلومة. وبناء عليه، فمن المحبذ استكمال تقييم الأثر بإجراء تحليل فعالية التكاليف أو تحليل المنافع والتكاليف.

دراسة حالة NUSAF: زيادة أهمية تقييم الأثر

التصميم الشامل

بالإضافة إلى تقييم مجمل آثار برنامج توفير الفرص للشباب، تمت الاستفادة من تقييم الأثر لاختبار تدخل تكميلي تجريبي على تغير تصميم مبتكر.. كانت الأدلة السردية من جولات سابقة من تمويل البرنامج تبين أن جودة الإدارة والتخطيط وخدمات الإرشاد التي يوفرها ميسرو المقاطعة والمجتمع المحلي هي محددات رئيسية لنجاح مجموعة الشباب. لذا، أراد تقييم الأثر التركيز على فاعلية تخصيص دفعة إضافية لتوظيف مستشار متابعة وإرشاد (MEA) من اختيار مجموعة الشباب نفسها.

وتم تعيين مجموعات المعالجة بشكل عشوائي ليشاركوا في التصميم الشامل. كما تم تعيين المشاريع التي سيتم تمويلها بشكل عشوائي لمجموعة من ثلاثة. فتمت معاملة مجموعة المعالجة الأولى بشكل طبيعي من دون تدخل إضافي. أما بالنسبة للمجموعة الثانية فكان موظفو المقاطعة يقيمون مستشاري المتابعة والإرشاد، وأما بالنسبة للمجموعة الثالثة، فتم توفير موارد إضافية وطلب من المجموعة أن تقيم مستشاري المتابعة والإرشاد بنفسها.



(...)

دراسة حالة NUSAF: زيادة أهمية تقييم الأثر

الأساليب المختلطة

استفاد تقييم برنامج توفير الفرص للشباب من الأسئلة الكمية والنوعية. إذ تم تقديم الاستبيان الكمي لما يقرب من ٢,٦٠٠ من الشباب في حين تم تقديم الاستبيان النوعي إلى ما يقرب من ١٠٠ من الشباب. وتضمنت الأسئلة النوعية الفئات موضع الاهتمام التالية:

١. جودة التفاعل والتعاون داخل المجموعة ، بما في ذلك تشكيل المجموعة وقيادة المجموعة وهيكلها وعمليات اتخاذ القرارات داخل المجموعة وأنشطة المجموعة السابقة والحاضرة والمستقبلية ومنافع وتحديات العمل ضمن مجموعات والأسباب الفردية لاختيار العمل ضمن مجموعة بالرغم من التحديات.
٢. تخصيص منح NUSAF، بما في ذلك أسلوب المجموعة لتقسيم المنح، وأولويات المجموعة في التمويل مقابل الخطط الأصلية الموضوعة للمشروع والانحراف والاستعمالات الأخرى غير الرسمية للمنح.
٣. الخبرة الناتجة عن التدريب، بما في ذلك اختيار تدريب على المهارات ضمن مجموعة، الثقة في القدرة على تطبيق المهارات المكتسبة، ومنافع وتحديات تطبيق المهارات كإستراتيجية لسبل العيش.
٤. استراتيجيات سبل العيش، بما في ذلك بناء سبل العيش بعد التدريب المهني؛ المخاطر والنجاح والفشل المرتبطة باستراتيجيات سبل العيش الجديدة؛ أسباب النجاح أو الفشل؛ استراتيجيات سبل العيش البديلة؛ واستراتيجيات أخرى للتعامل مع المخاطر والصدمات.
٥. تمكين المجتمع المحلي ومشاركته، بما في ذلك حس الانتماء إلى المجتمعات المحلية والمشاركة المدنية والعلاقات بين الجنسين والدعم الاجتماعي والعوائق الاجتماعية والعلاقات مع الجيران.

لا يمكن للبيانات الكمية وحدها أن تبين غنى البرنامج بأكمله. لذا، تستخدم الإجابات عن هذه الأسئلة لتحقيق فهم أفضل لكيفية تغيير NUSAF لحياة المشاركين، وكذلك لتقديم بعض القصص التي ستساعد في فهم كيفية تأثير البرنامج في حياة الأفراد.

المصدر: باتمان ومارتينيز وفيالا (٢٠٠٩).

قراءات رئيسية

- Bamberger, M., Rao, V., and Woolcock, M. 2010. *Using Mixed Methods in Monitoring and Evaluation: Experiences from International Development*. Policy Research Working Paper 5245, Washington, DC: The World Bank.
http://www-wds.worldbank.org/servlet/WDSCContentServer/WDSP/IB/2010/03/23/000158349_20100323100628/Rendered/PDF/WPS5245.pdf
- Bamberger, M., Rugh, J., and Mabry, L. 2006. *Real World Evaluation: Working under Budget, Time, Data and Political Constraints*. Thousand Oaks: Sage Publications. (See chapter 13.) <http://realworldevaluation.org/>
- Cellini, S. R., and Kee, J. E. 2010. "Cost-Effectiveness and Cost-Benefit Analysis." In: Wholey, J., Hatry, H. P., and Newcomer, K. E., eds. *Handbook of Practical Program Evaluation*, 3rd ed. San Francisco: Jossey-Bass.
<http://home.gwu.edu/~scellini/CelliniKee21.pdf>
- Knowles, J., and Behrman, J. 2005. *A Practical Guide to Economic Analysis of Youth Projects*. HNP Discussion Paper. Washington, DC: The World Bank.
<http://siteresources.worldbank.org/HEALTHNUTRITIONANDPOPULATION/Resources/281627-1095698140167/KnowlesPracticalGuideFinal.pdf>

بنود

- Almeida, R., and Galasso, E. 2008. *Jumpstarting self-employment? Evidence from welfare participants in Argentina*. Washington, DC: The World Bank. <http://ftp.iza.org/dp2902.pdf>
- Asian Development Bank. 2007. *Handbook on Social Analysis, A Working Document*. Manila: ADB. <http://www.adb.org/Documents/Handbooks/social-analysis/Appendixes.pdf>
- Attanasio, O., Kugler, A., and Meghir, C. 2009. "Subsidizing Vocational Training for Disadvantaged Youth in Developing Countries: Evidence from a Randomized Trial." IZA Discussion Paper No. 4251. Bonn: IZA. <http://ftp.iza.org/dp4251.pdf>
- Baker, J. 2000. *Evaluating the Impact of Development Projects on Poverty: A Handbook for Practitioners*. Washington, DC: The World Bank. <http://siteresources.worldbank.org/INTISPMA/Resources/handbook.pdf>
- Bamberger, M., Rao, V., and Woolcock, M. 2010. "Using Mixed Methods in Monitoring and Evaluation: Experiences from International Development." Policy Research Working Paper No. 5245. Washington, DC: The World Bank. http://www-wds.worldbank.org/servlet/WDSContentServer/WDSP/IB/2010/03/23/000158349_20100323100628/Rendered/PDF/WPS5245.pdf
- Beauvy-Sany, M., Conklin, S., Hershkowitz, A., Rajkotia, R., and Berg, C. 2009. *Guidelines and Experiences for Including Youth in Market Assessments for Stronger Youth Workforce Development Programs*. Washington, DC: The SEEP Network. http://www.seepnetwork.org/Resources/YouthPLP_Assessments.pdf
- Bertrand, A., Beauvy-Sany, M., Cilimkovic, S., Conklin, S., and Jahic, S. 2009. *Monitoring and Evaluation for Youth Workforce Development Projects*. Washington DC: The SEEP Network. http://seepnetwork.org/Resources/YouthPLP_MonitoringEval.pdf
- Betcherman, G., Godfrey, M., Puerto, S., Rother, F., and Stavreska, A. 2007. "A Review of Interventions to Support Young Workers: Findings of the Youth Employment Inventory." SP Discussion Paper No. 0715. Washington, DC: The World Bank. <http://www.youth-employment-inventory.org/downloads/1.pdf>
- Blattman, C., Fiala, N., and Martinez, S. 2011. *Employment Generation in Rural Africa: Mid-term Results from an Experimental Evaluation of Youth Opportunities Program in Northern Uganda*. New Haven, CT: Innovations for Poverty Action. <https://www.poverty-action.org/sites/default/files/blattmanfialamartinez.midtermreport.pdf>
- Bose, R. 2010. "A Checklist for the Reporting of Randomized Control Trials of Social and Economic Policy Interventions in Developing countries: CEDE Version 1.0." Working Paper No. 6. New Delhi: International Initiative for Impact Evaluation. http://www.3ieimpact.org/admin/pdfs_papers/working paper 6.pdf

- Brady, M., Salem, A., and Zibani, N. 2007. "Bringing New Opportunities to Adolescent Girls in Socially Conservative settings: The Ishraq Program in Rural Upper Egypt." *Transitions to Adulthood*, Brief No. 12. New York: Population Council. <http://www.popcouncil.org/pdfs/IshraqFullReport.pdf>
- Brener, N. D., Billy, J. O. G., and Grady, W.R. 2003. "Assessment of Factors Affecting the Validity of Self-Reported Health-Risk Behavior Among Adolescents: Evidence From the Scientific Literature." *Journal of Adolescent Health* 33: 436–457. <http://www.cdc.gov/HealthyYouth/yrbs/pdf/validity.pdf>
- Bronfenbrenner, U. 1979. *The Ecology of Human Development: Experiments by Nature and Design*. Cambridge, MA: Harvard University Press. http://books.google.com/books?id=OCmbzWka6xUC&printsec=frontcover&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false
- Bruhn, M., and Zia, B. 2011. "Stimulating Managerial Capital in Emerging Markets: The Impact of Business and Financial Literacy for Young Entrepreneurs." Policy Research Working Paper No. 5642. Washington, DC: The World Bank. http://www-wds.worldbank.org/external/default/WDSCContentServer/WDSP/IB/2011/04/27/000158349_20110427082512/Rendered/PDF/WPS5642.pdf
- Card, D., Kluve, J., and Weber, R. 2009. "Active Labor Market Policy Evaluations: A Meta-Analysis." IZA Discussion Paper No. 4002. Bonn. <http://ftp.iza.org/dp4002.pdf>
- Cooley, L. 1989. "The Logical Framework: Program Design for Program Results." *The Entrepreneurial Economy Review* 8(1): 8–15. <http://evaluation.zunia.org/post/the-logical-framework-program-design-for-program-results/>
- Creswell, John. 2008. *Research Design: Qualitative, Quantitative and Mixed Methods Approaches*. Thousand Oaks, CA: Sage. http://books.google.com/books?id=bttwENORfhgC&printsec=frontcover&source=gbs_ge_summary_r&cad=0
- Cunningham, W., Sanchez-Puerta, M. L., and Wuermli, A. 2010. "Active Labor Market Policies for Youth: A Framework to Guide Youth Employment Interventions." Employment Policy Primer No. 16. Washington, DC: The World Bank. http://siteresources.worldbank.org/INTLTM/214578-1103128720951/22795057/EPPNoteNo16_Eng.pdf
- Davies, R. J., and Dart, J. 2005. *The "Most Significant Change" (MSC) Technique: A Guide to Its Use*. <http://mande.co.uk/docs/MSCGuide.pdf>
- Development Marketplace. 2008. *Development Marketplace Grantee Toolkit: Project Design, Monitoring and Evaluation for Small Innovative Projects*. Washington DC: The World Bank. http://siteresources.worldbank.org/DEVMARKETPLACE/Resources/DM_Grantee_Toolkit_Final.pdf

- DFID (Department for International Development). 1999. "Framework." Sustainable Livelihood Guidance Sheets, Section 2. Oxford: DFID. <http://www.eldis.org/vfile/upload/1/document/0901/section2.pdf>
- Duflo, E., Glennerster, R., and Kremer, M. 2006. "Using Randomization in Development Economics Research: A Toolkit." BREAD Working Paper No. 136. <http://www.povertyactionlab.org/sites/default/files/documents/Using%20Randomization%20in%20Development%20Economics.pdf>
- Fiszbein, A., and Schady, N. 2009. *Conditional Cash Transfer, Reducing Present and Future Poverty*. Policy Research Report. Washington, DC: The World Bank. http://siteresources.worldbank.org/INTCCT/Resources/5757608-1234228266004/PRR-CCT_web_noembargo.pdf
- GAO (U.S. Government Accountability Office). 1991. *Designing Evaluations*. Washington, DC: GAO. http://www.gao.gov/special.pubs/10_1_4.pdf
- Gawler, M. 2005. *Useful Tools for Engaging Young People in Participatory Evaluation*. Geneva: UNICEF CEE/CIS. <http://www.artemis-services.com/downloads/tools-for-participatory-evaluation.pdf>
- Gertler, P., Martinez, S., Premand, P., Rawlings, L., and Vermeersch, C. 2011. *Impact Evaluation in Practice*. Washington, DC: The World Bank. <http://www.worldbank.org/ieinpractice>
- Gosparini, P., Russo, L., Sirtori, M., and Valmarana, C. 2004. *The Monitoring and Evaluation Manual of the NGOs of the Forum Solint*. Rome, Italy: CISP, COSV, COOPI, Intersos, Movimondo, and DRN. http://www.cosv.org/echoTrain/materiale/0B_ITA/ECHOTrain_Documenti/ECHOTrain_Documenti_Manuali/ECHOTrain_Documenti_Manuali_SOLINT/Manuale M&E- Solint.pdf
- Hempel, K. 2006. *Erfolgskontrolle in der deutschen Entwicklungszusammenarbeit: Zentrale Herausforderungen unter besonderer Berücksichtigung der Transparenz gegenüber der Öffentlichkeit*. Stuttgart: ibidem-Verlag.
- Ibarrarán, P., and Rosas Shady, D. 2009. "Evaluating the Impact of Job Training Programmes in Latin America: Evidence from IDB-Funded Operations." *Journal of Development Effectiveness* 1(2): 195–216. http://www.iza.org/conference_files/ELMPDC2009/ibarraran_p4263.pdf
- Imas, L., and Rist, R. 2009. *The Road to Results: Designing and Conducting Effective Development Evaluations*. Washington, DC: The World Bank. http://books.google.com/books?id=NEsg-BtinIsC&printsec=frontcover&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false

Innovations for Poverty Action. 2010. *IPA IRB Submission Guidelines and Application*. Internal Document. New Haven, CT: IPA.

Jaramillo, M., and Parodi, S. 2003. Jóvenes Emprendedores. Lima, Peru: Instituto Apoyo. <http://www.grade.org.pe/download/pubs/MJ-SP-J%C3%B3venes%20emprendedores.pdf>

Jones, N., Jones, H., Steer, L., and Datta, A. 2009. "Improving Impact Evaluation Production and Use." Working Paper No. 3000. London: Overseas Development Institute. <http://www.odi.org.uk/resources/download/3177.pdf>

J-PAL (Abdul Latif Jameel Poverty Action Lab). 2011. "Cost-benefit/Effectiveness/Comparison Analysis." J-PAL Web site. <http://www.povertyactionlab.org/methodology/what-evaluation/cost-benefit/effectiveness/comparison-analyses>. Accessed May 31, 2011.

Karlan, D. 2009. "Thoughts on Randomised Trials for Evaluation of Development: Presentation to the Cairo Evaluation Clinic." *The Journal of Development Effectiveness* 1(3): 237–242. http://www.3ieimpact.org/admin/pdfs_papers/S0.pdf

Kellogg (W. K. Kellogg Foundation). 1998. *Evaluation Handbook*. Battle Creek, MI: W. K. Kellogg Foundation. <http://www.wkkf.org/knowledge-center/resources/2010/W-K-Kellogg-Foundation-Evaluation-Handbook.aspx>

———. 2004. *Logic Model Development Guide*. Battle Creek, MI: W. K. Kellogg Foundation. <http://www.wkkf.org/knowledge-center/resources/2006/02/WK-Kellogg-Foundation-Logic-Model-Development-Guide.aspx>

Khandker, S., Koolwal, G., and Samad, H. 2010. *Handbook on Impact Evaluation: Quantitative Methods and Practices*. Washington, DC: The World Bank. http://www-wds.worldbank.org/external/default/WDSCContentServer/IW3P/IB/2009/12/10/000333037_20091210014322/Rendered/PDF/S20990PUB0EPI1101Official0Use0Only1.pdf

Klinger, B., and Schuendeln, M. 2007. "Can Entrepreneurial Activity be Taught? Quasi-Experimental Evidence from Central America." Working Paper No. 153. Cambridge, MA: Harvard University Center for International Development. http://www.hks.harvard.edu/var/ezp_site/storage/fckeditor/file/pdfs/centers-programs/centers/cid/publications/faculty/wp/153.pdf

Kusek, J. Z., and Rist, R. C. 2004. *Ten Steps to a Results Based Monitoring and Evaluation System: A Handbook for Development Practitioners*. Washington, DC: The World Bank. http://www-wds.worldbank.org/external/default/WDSCContentServer/WDSP/IB/2004/08/27/000160016_20040827154900/Rendered/PDF/296720PAPER0100steps.pdf

- Leeuw, F., and Vaessen, J. 2009. *Impact Evaluations and Development: NONIE Guidance on Impact Evaluation*. Washington, DC: Network of Networks on Impact Evaluation. http://siteresources.worldbank.org/EXTOED/Resources/nonie_guidance.pdf
- Muzi, S. 2011. "Summary Notes on Impact Evaluations in the Active Labor Market/Youth Employment Programs Cluster Supported by SIEF." Washington, DC: The World Bank.
- NIH (National Institutes of Health). 2008. *Protecting Human Research Participants*. Bethesda: NIH Office of Extramural Research. <http://phrp.nihtraining.com/users/PHRP.pdf>
- NORC (National Opinion Research Center). 2007. *Operational Plan for the Baseline Survey of "Mi Primer Empleo" Youth Employment Program, Honduras*. Chicago: NORC.
- OECD (Organization for Economic Cooperation and Development). 1991. *Principles for Evaluation of Development Assistance*. Paris: Development Assistance Committee. <http://www.oecd.org/dataoecd/31/12/2755284.pdf>
- . (n.d.). DAC Criteria for Evaluating Development Assistance. http://www.oecd.org/document/22/0,2340,en_2649_34435_2086550_1_1_1_1,00.html
- Osborn, D., and Gaebler, T. 1992. *Reinventing Government: How the Entrepreneurial Spirit is Transforming the Public Sector*. Reading, MA: Addison-Wesley.
- Penrose-Buckley, C. 2007. "Annex 2: Rapid Assessment of Markets and Producers." In: *Producer Organisations: A Guide to Developing Collective Rural Enterprises*. Oxford: Oxfam GB. http://www.oxfam.org.uk/resources/downloads/produorgs_book.pdf
- Powers, J. L., and Tiffany, J. S. 2006. "Engaging Youth in Participatory Research and Evaluation." *Journal of Public Health Management and Practice*, November (Suppl.), S79-S87. http://www.health.state.ny.us/community/youth/development/docs/jphmp_s079-s087.pdf
- Ravallion, M. 2008. *Evaluating Anti-Poverty Programs. Handbook of Development Economics*, Vol. 4, Schultz, T. P., and Strauss, J. (eds). Amsterdam: Elsevier. http://siteresources.worldbank.org/INTISPMA/Resources/383704-1130267506458/Evaluating_Antipoverty_Programs.pdf
- Rawlings, L., and Rubio, G. 2005. "Evaluating the Impacts of Conditional Cash Transfer Programs." *World Bank Research Observer* 20(1): 29-55. http://www.crin.org/docs/Evaluating_the_Impact_of_Cash_Transfer_Programs.pdf
- RMC (RMC Research Corporation). 2009. *"Our Nation" Evaluation*. Denver, CO: RMC. http://www.myja.org/programs/evaluation/reports/our_nation.pdf

- Roberts, B., Caspi, A., and Moffitt, T. 2003. "Work Experiences and Personality Development in Young Adulthood." *Journal of Personality and Social Psychology* 84(3): 582–593. <http://www.mendeley.com/research/work-experiences-and-personality-development-in-young-adulthood-1/>
- Robins, R., Fraley, R., Roberts, B., and Trzesniewski, K. 2001. "A Longitudinal Study of Personality Change in Young Adulthood." *Journal of Personality* 69(4): 617–640. <http://internal.psychology.illinois.edu/~broberts/Robins,%20et%20al,%202001.pdf>
- Rubio, G. 2011. "The Design and Implementation of a Menu of Evaluations." PREM Notes No. 6, The Nuts and Bolts of M and E Systems. Washington, DC: The World Bank. <http://siteresources.worldbank.org/INTPOVERTY/Resources/335642-1276521901256/premnoteME6.pdf>
- Rossiasco, P., O'Neil, F., Richardson, A., and Francis, P. 2010. "Youth and Employment In Post-Conflict Countries: The Psycho-Social Dimension." *Child and Youth Development Notes* (4)2. Washington DC: The World Bank. http://siteresources.worldbank.org/INTLM/Resources/390041-1319047943696/CYDN_No2_Psychosocial_Employment.pdf
- Sabo Flores, K. 2008. *Youth Participatory Evaluation: Strategies for Engaging Young People*. San Francisco: Jossey-Bass. <http://www.josseybass.com/WileyCDA/WileyTitle/productCd-0787983926.html>
- Savedoff, W. D., Levine, R., and Birdsall, N. 2006. *When Will We Ever Learn? Improving Lives through Impact Evaluation*. Washington, DC: Center for Global Development. <http://www.cgdev.org/content/publications/detail/7973>
- Shapiro, J. 2003. "Monitoring and Evaluation." Washington, DC: CIVICUS. [http://www.civicus.org/new/media/Monitoring and Evaluation.doc](http://www.civicus.org/new/media/Monitoring%20and%20Evaluation.doc)
- Taylor-Powell, E. 2005. "Logic Models: A Framework for Program Planning and Evaluation." University of Wisconsin–Extension, Program Development and Evaluation, 2005. slides 15–16. <http://www.uwex.edu/ces/pdande/evaluation/pdf/nutritionconf05.pdf>. Accessed June 2, 2011.
- UNICEF (United Nations Children's Fund). 1991. "A UNICEF Guide for Monitoring and Evaluation: Making a Difference?" New York: UNICEF. <http://preval.org/documentos/00473.pdf>
- Trochim, W. M. K. 2006. "The Nonquivalent Groups Design." Web Center for Social Research Methods, Research Methods Knowledge Base. <http://www.socialresearch-methods.net/kb/quasnegd.php>. Accessed June 2, 2011.
- World Bank. 2002. *A Sourcebook for Poverty Reduction Strategies*. Washington, DC: The World Bank. http://siteresources.worldbank.org/INTPRS1/Resources/383606-1205334112622/4943_annex_c.pdf

- . 2006. *World Development Report 2007: Development and the Next Generation*. Washington, DC: The World Bank. http://www-wds.worldbank.org/external/default/WDSCContentServer/WDSP/IB/2006/09/13/000112742_20060913111024/Rendered/PDF/359990WDR0complete.pdf
- . 2007a. *Data for Impact Evaluation*. Doing Impact Evaluation No. 6. Washington, DC: The World Bank. http://siteresources.worldbank.org/INTISPMA/Resources/383704-1146752240884/Doing_ie_series_06.pdf
- . 2007b. *Impact Evaluation for Microfinance: Review of Methodological Issues*. Doing Impact Evaluation No. 7. Washington, DC: The World Bank. http://siteresources.worldbank.org/INTISPMA/Resources/383704-1146752240884/Doing_ie_series_07.pdf
- . 2007c. "Miles to Go: A Quest for an Operational Labor Market Paradigm for Developing Countries." Washington, DC: The World Bank. http://siteresources.worldbank.org/INTLM/Resources/390041-1212776476091/5078455-1267646113835/MILESThequestoperationalLmparadigm_Jan212008.pdf
- . 2008. Project Paper for a project in the amount of USD 2,793,706 to the Republic of Liberia for a proposed Economic Empowerment of Adolescent Girls Project, Report No. 44985-LR. Washington, DC: The World Bank.
- . 2009. *Institutionalizing Impact Evaluation within the Framework of a Monitoring and Evaluation System*. Washington, DC: The World Bank. http://siteresources.worldbank.org/EXTEVACAPDEV/Resources/4585672-1251461875432/inst_ie_framework_me.pdf
- . 2011. "Evaluation of the Use of Entertainment Education to Improve Financial Capability in South Africa." Concept Note (February). Washington, DC: The World Bank.

مؤلفات خاصة بتنمية/توظيف الشباب

- Bidwell, K., Galbraith, C., et al. 2008. *Market Assessment Toolkit for Vocational Training Providers and Youth*. New York: Women's Commission for Refugee Women and Children and Columbia University School of International and Public Affairs.
http://www.womensrefugeecommission.org/docs/ug_ysl_toolkit.pdf
- Bronfenbrenner, U. 1979. *The Ecology of Human Development: Experiments by Nature and Design*. Cambridge, MA: Harvard University Press.
http://books.google.com/books?id=OCmbzWka6xUC&printsec=frontcover&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false
- Card, D., Ibararán, P., and Villa, J. M. 2011. "Building in an Evaluation Component for Active Labor Market Programs: A Practitioner's Guide." IZA Discussion Paper No. 6085. Bonn: IZA. <http://ftp.iza.org/dp6085.pdf>
- Cunningham, W., Cohan, L., Naudeau, S., and McGinnis, L. 2008. *Supporting Youth at Risk: A Policy Toolkit for Middle-Income Countries*. Washington, DC: The World Bank.
<http://siteresources.worldbank.org/INTCY/Resources/395766-1187899515414/SupportingYouthAtRisk.pdf>
- Cunningham, W., Sanchez-Puerta, M. L., and Wuermli, A. 2010. "Active Labor Market Policies for Youth: A Framework to Guide Youth Employment Interventions." Employment Policy Primer No. 16. Washington, DC: The World Bank.
http://siteresources.worldbank.org/INTLM/214578-1103128720951/22795057/EPPNoteNo16_Eng.pdf
- International Labour Office. 2010. *Global Employment Trends for Youth: Special Issues on the Impact of the Global Economic Crisis on Youth*. Geneva: ILO.
http://www.ilo.org/wcmsp5/groups/public/---ed_emp/---emp_elm/---trends/documents/publication/wcms_143349.pdf
- James-Wilson, D. 2008. *Youth Livelihoods Development Program Guide*. Washington DC: EQUIP3. <http://www.equip123.net/docs/e3-LivelihoodsGuide.pdf>
- Making Cents International. *State of the Field in Youth Enterprise Development and Livelihoods Development series*. Washington, DC: MCI.
<http://www.youtheconomicopportunities.org/media.asp>
- Van Adams, A. 2007. *The Role of Youth Skills in the Transition to Work: A Global Review*. Washington, DC: The World Bank.
<http://siteresources.worldbank.org/INTCY/Resources/395766-1187899515414/RoleofYouthSkills.pdf>

Van der Geest, K. 2010. *Rural Youth Employment in Developing Countries: A Global View*. Rome: FAO. <http://www.fao.org/docrep/012/al414e/al414e00.pdf>

Women's Refugee Commission. 2009. *Building Livelihoods: A Field Manual for Practitioners in Humanitarian Settings*. New York: WRC. http://www.womensrefugeecommission.org/docs/livelihoods_manual.pdf

World Bank 2006. *World Development Report 2007: Development and the Next Generation*. Washington, DC: The World Bank. http://www-wds.worldbank.org/external/default/WDSCContentServer/WDSP/IB/2006/09/13/000112742_20060913111024/Rendered/PDF/359990WDR0complete.pdf

مؤلفات خاصة بالمتابعة والتقييم

Bamberger, M., Rugh, J., and Mabry, L. 2006. *Real World Evaluation: Working under Budget, Time, Data and Political Constraints*. Thousand Oaks: Sage. <http://www.realworldevaluation.org>

Bertrand, A., Beauvy-Sany, M., Cilimkovic, S., Conklin, S., and Jahic, S. 2009. *Monitoring and Evaluation for Youth Workforce Development Projects*. Washington DC: The SEEP Network. http://seepnetwork.org/Resources/YouthPLP_MonitoringEval.pdf

Catley, A., Burns, J., Abebe, D., and Suji, O. 2010. *Participatory Impact Assessment, A Guide for Practitioners*. Medford, MA: Feinstein International Center. https://wikis.uit.tufts.edu/confluence/download/attachments/19924843/Part_Impact_10_21_08V2.pdf?version=1&modificationDate=1225200269000

Duflo, E., Glennerster, R., and Kremer, M. 2006. "Using Randomization in Development Economics Research: A Toolkit." BREAD Working Paper No. 136. <http://www.povertyactionlab.org/sites/default/files/documents/Using%20Randomization%20in%20Development%20Economics.pdf>

Gertler, P. Martinez, S., Premand, P., Rawlings, L., and Vermeersch, C. 2011. *Impact Evaluation in Practice*. Washington, DC: The World Bank. <http://www.worldbank.org/ieinpractice>

Imas, L., and Rist, R. 2009. *The Road to Results: Designing and Conducting Effective Development Evaluations*. Washington, DC: The World Bank. http://books.google.com/books?id=NEsg-BtinIsC&printsec=frontcover&source=gbs_ge_summary_r&cad=0

Karlan, D., and Appel, J. 2011. *More than Good Intentions: How a New Economics Is Helping to Solve Global Poverty*. Boston, MA: Dutton. http://books.google.com/books?id=JOMDsm5Gn9kC&printsec=frontcover&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false

- Khandker, S., Koolwal, G., and Samad, H. 2010. *Handbook on Impact Evaluation: Quantitative Methods and Practices*. Washington, DC: The World Bank.
http://www-wds.worldbank.org/external/default/WDSCContentServer/IW3P/IB/2009/12/10/000333037_20091210014322/Rendered/PDF/S20990PUB0EPI1101Official0Use0Only1.pdf
- Kluve, J. 2011. *Measuring Employment Effects of Technical Cooperation Interventions: Some Methodological Guidelines*. Second revised edition. Eschborn: GIZ
<http://www2.gtzt.de/wbf/4tDx9kw63gma/Methodenleitfaden.pdf>
- Knowles, J., and Behrman, J. 2005. *A Practical Guide to Economic Analysis of Youth Projects*. HNP Discussion Paper. Washington, DC: The World Bank.
<http://siteresources.worldbank.org/HEALTHNUTRITIONANDPOPULATION/Resources/281627-1095698140167/KnowlesPracticalGuideFinal.pdf>
- Kusek, J. Z., and Rist, R. C. 2004. *Ten Steps to a Results-Based Monitoring and Evaluation System: A Handbook for Development Practitioners*. Washington, DC: The World Bank.
<http://www.oecd.org/dataoecd/23/27/35281194.pdf>
- Leeuw, F., and Vaessen, J. 2009. *Impact Evaluations and Development: NONIE Guidance on Impact Evaluation*. Washington, DC: Network of Networks on Impact Evaluation.
http://siteresources.worldbank.org/EXTOED/Resources/nonie_guidance.pdf
- Powell, E., Jones, L., and Henert, E. 2003. *Enhancing Program Performance with Logic Models*. University of Wisconsin–Extension, Program Development and Evaluation.
<http://www1.uwex.edu/ces/lmcourse/>
- Ravallion, M. 2008. *Evaluating Anti-Poverty Programs*. *Handbook of Development Economics*, Vol. 4, Schultz, T. P., and Strauss, J. (eds). Amsterdam: Elsevier.
http://siteresources.worldbank.org/INTISPMA/Resources/383704-1130267506458/Evaluating_Antipoverty_Programs.pdf
- Wholey, J., Hatry, H. P., and Newcomer, K. E., eds. 2010. *Handbook of Practical Program Evaluation*, 3rd ed. San Francisco: Jossey-Bass.
http://books.google.com/books?id=zntNhoO6gCUC&printsec=frontcover&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false
- World Bank 2007. *Doing Impact Evaluation*. A series of fourteen associated reports.
<http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTPOVERTY/EXTISPMA/0,,contentMDK:22268536~pagePK:210058~piPK:210062~theSitePK:384329,00.html>

قواعد بيانات بحثية

EconPapers <http://econpapers.repec.org/>

Education Resources Information Center (ERIC) <http://www.eric.ed.gov/>

IDEAS <http://ideas.repec.org/>

IZA Discussion Papers <http://www.iza.org/en/webcontent/publications/papers>

J-STOR <http://www.jstor.org/>

National Bureau of Economic Research (NBER) <http://www.nber.org/>

PsychInfo <http://www.apa.org/pubs/databases/psycinfo/index.aspx>

Social Science Research Network
<http://papers.ssrn.com/sol3/DisplayAbstractSearch.cfm>

الدوريات الأكاديمية

Journal of Adolescence
http://www.elsevier.com/wps/find/journaldescription.cws_home/622849/description#description

Journal of Early Adolescence
<http://jea.sagepub.com/>

Journal of Research on Adolescence
<http://www.wiley.com/bw/journal.asp?ref=1050-8392>

Journal of Youth and Adolescence
<http://www.springer.com/psychology/child+%26+school+psychology/journal/10964>

Journal of Youth Development
www.nae4ha.org/directory/jyd/index.html

Journal of Youth Studies
<http://taylorandfrancis.co.uk/journals/titles/13676261.asp>

New Directions for Child and Adolescent Development
<http://www3.interscience.wiley.com/journal/85511342/home>

New Directions for Youth Development
[http://onlinelibrary.wiley.com/journal/10.1002/\(ISSN\)1537-5781/homepage/ProductInformation.html](http://onlinelibrary.wiley.com/journal/10.1002/(ISSN)1537-5781/homepage/ProductInformation.html)

The Future of Children

<http://futureofchildren.org/>

Vulnerable Children and Youth Studies

<http://www.informaworld.com/smpp/title~db=all~content=t724921266>

Young: Nordic Journal of Youth Research

<http://you.sagepub.com>

قواعد بيانات تقييمات الأثر القائمة والجارية

3ie International Initiative for Impact Evaluation

http://www.3ieimpact.org/database_of_impact_evaluations.html

Abdul Latif Jameel Poverty Action Lab (JPAL)

http://www.povertyactionlab.org/search/apachesolr_search?filters=type:evaluation

Innovations for Poverty Action (IPA)

<http://www.poverty-action.org/project-evaluations/search>

World Bank DIME Initiative

<http://web.worldbank.org/WBSITE/EXTERNAL/EXTDEC/EXTDEVIMPEVAINI/0,,contentMDK:21553788~pagePK:64168445~piPK:64168309~theSitePK:3998212,00.html>

Youth Employment Inventory

<http://www.youth-employment-inventory.org/>

المدونات

Aid Watch

<http://aidwatchers.com/>

Chris Blattman

<http://chrisblattman.com/>

Development that Works (IDB)

http://blogs.cross-cutting.org/desarrolloefectivo_en/

Find What Works

<http://findwhatworks.wordpress.com/>

Good Intentions Are Not Enough

<http://goodintentions.org/>

Innovations for Poverty Action

<http://www.poverty-action.org/blog>

Larry Dershem, Save the Children

<http://designmonitoringevaluation.blogspot.com/>

Monitoring and Evaluation News

<http://mande.co.uk/>

World Bank Development Impact

<http://blogs.worldbank.org/impactevaluations/>

بناء القدرات

ILO International Training Centre Course on M&E for Youth Employment Projects

<http://www.itcilo.org>

International Program for Development Evaluation Training (IPDET)

<http://www.ipdet.org>

J-PAL Course on Impact Evaluation

<http://www.povertyactionlab.org/course>

World Bank Impact Evaluation Workshops

<http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTPOVERTY/EXTISPMA/0,,contentMDK:21754074~menuPK:384336~pagePK:148956~piPK:216618~theSitePK:384329,00.html>

YEN Evaluation Clinics

<http://yenclinic.groupsite.org>

يمكنك استعراض مجموعة واسعة من المراجع الخاصة بالتعلم الذاتي عبر شبكة الانترنت والتدريبات الميدانية على المتابعة والتقييم (غير تقييم الأثر) على موقع My M&E على شبكة الانترنت

http://www.mymande.org/index.php?q=training_search&x=admin

المواقع الإلكترونية

Evalsed: The resource for the evaluation of socioeconomic development

http://ec.europa.eu/regional_policy/sources/docgener/evaluation/evalsed/

My M&E

<http://www.mymande.org/>

Web Center for Social Research Methods

<http://www.socialresearchmethods.net/kb/design.php>

ملحق ١: نموذج مؤشرات لتقييم الشباب

حالة الشباب		
المؤشرات المختارة	الفترة	
• عدد ونسبة الشباب الذين يعيشون على أقل من ٢ دولار يومياً • عدد ونسبة الشباب الذين يعانون من سوء التغذية • عدد ونسبة الشباب الذين يعيشون في مأوى بدون البنية التحتية الأساسية	الفقر	
• معدلات معرفة القراءة والكتابة لدى الشباب • صافي الالتحاق بالتعليم الابتدائي • صافي الالتحاق بالتعليم الثانوي • إجمالي الالتحاق بالتعليم بعد الثانوي • نسبة التحاق الإناث إلى الذكور في مختلف مستويات التعليم • التحصيل العلمي في عمر (سن) X	المستوى التعليمي	
• نسبة الشباب العاملين والعاطلين جزئياً عن العمل • معدل مشاركة الشباب في القوى العاملة • نسبة البطالة الشباب إلى الكبار • عدد ونسبة الشباب غير المتحقين بالمدرسة أو غير المتحقين بتدريب أو العاطلين عن العمل • توظيف الشباب حسب القطاع (أو حسب نوع الوظيفة) • توظيف الشباب حسب نوع العمل (الأجر، العمل لحسابه الخاص، رب عمل، عمل عائلي وعمل غير مدفوع) • توظيف الشباب حسب نوع العقد (عقد رسمي أو عقد غير رسمي، دوام كامل أو دوام جزئي، غير محدد المدة أو مؤقت) • متوسط الدخل الشهري • متوسط الوقت الذي يستغرقه العثور على وظيفة • توفر البرامج لتدريب المياني والتدريب المهني • تفضيلات الوظيفة للشباب (حكومية، قطاع خاص، العمل لحسابه الخاص، وغيرها) واتجاهاتهم نحو العمل	الحالة الوظيفية	
• انتشار داء نقص المناعة المكتسبة (الإيدز) بين الشباب (ب.٪) • عدد ونسبة الشباب الذين لديهم معرفة صحيحة ببدء نقص المناعة المكتسبة (الإيدز) ومنع الحمل • عدد ونسبة الشباب الذين يستخدمون وسائل منع الحمل في الوقت الحالي • عدد ونسبة الشباب الذين يأنغوا عن ممارسة الجنس المحفوف بالمخاطر في العام الماضي • انتشار تعاطي المخدرات والكوكايين والموراف المستنشقة) • انتشار استهلاك التبغ والكحول بين الشباب	الصحة	
• الوسيط لسن بدء النشاط الجنسي • عدد ونسبة النساء في سن ١٥-١٩ الذين لديهن أطفال أو حوامل • حصّة السكان من الإناث ممن لديهم طفل واحد على الأقل، حسب السن • نسبة الفتحات المزروجات دون سن ١٨ سنة • الوسيط لسن الزواج الأول، حسب النوع (الجنس)	الوضع الأسري	
• عدد ونسبة الشباب مرتكبي أو ضحايا الجريمة والعنف • عدد ونسبة الشباب من أفراد العصابات أو الجماعات المتعردة • عدد ونسبة الشباب المتأثرين بالعنف المنزلي • عدد ونسبة الشباب المتأثرين بالعنف الجنسي • عدد ونسبة الشباب على اتصال بنظام العدالة	التاريخ الجنائي	

نموذج مؤشرات لتقييم الشباب

المواطنة	• عدد ونسبة الشباب في سن التصويت الذين يسمون في الانتخابات • عدد ونسبة الشباب الذين يشتركون/يظهرون في الأنشطة المجتمعية • عدد ونسبة الشباب المشاركين في المنظمات والمجالس الشبابية وغيرها. • أسباب عدم مشاركة الشباب في الأنشطة المدنية
مواطن الضعيف	• عدد ونسبة الشباب الذين يعيشون في الشارع • عدد ونسبة الشباب المتأثرين بالدعارة و/أو الاتجار في البشر • عدد ونسبة الشباب المتأثرين بالسرقات • عدد ونسبة الشباب من ذوي الإعاقات العقلية و/أو الجسمية • ما هي نقاط الضعف الخاصة بكل من الشباب والشابات؟
التصورات والتطلعات	• عدد ونسبة الشباب الذين يشعرون بالرضا عن حياتهم الحالية • عدد ونسبة الشباب الواثقين من المستقبل • عدد ونسبة الشباب الذين يرغبون في الهجرة • عدد ونسبة الشباب الذين لديهم ثقة في الكبار وقادة المجتمع والسياسيين والمؤسسات وما إلى ذلك. وبالنسبة لهؤلاء الذين لا يشعرون بالرضا، ماذا يرغبون أن يخبروهم؟ • ما التطلعات التي لدى الشباب لتعليمهم؟ وكيف يتصورون ويقارنون بين التعليم المهني والتعليم الجامعي ولماذا؟ • ما الذي يعبر عنه الشباب باعتباره أولويات وطموحات وفرص بالنسبة لهم؟ • ما هي التحديات والمخاطر والإجباطات الأكثر إلحاحا التي تواجه الشباب؟
البيئة المباشرة	
الفترة	المؤشرات المختارة
الأسرة	• متوسط حجم الأسرة • نسبة الأسر التي لديها ولد واحد • أنماط عمل الوالدين (الحالة الوظيفية وساعات العمل وغيرها) • متوسط عدد الساعات التي يقضيها الوالدين مع أفعالهم أسبوعياً • العلاقات بين الوالدين والأبناء
المدرسة	• عدد ونسبة الشباب في التعليم الثانوي في غضون X كم • أنواع الممارس ومؤسسات التدريب في المنطقة • جودة الممارس القائمة ومؤسسات التدريب في المنطقة • انتشار الرسوم المدرسية • تصورات المعلمين عن الشباب
الجوار	• الوضع الاجتماعي والاقتصادي وعدم المساواة • الديناميات الاجتماعية في داخل المجتمع (مثل: الثقة والتشبيك الاجتماعي وغيرها). • انتشار الجريمة والعنف • تفاعل الأقربان
الخدمات الملائمة للشباب	• عدد الخدمات الصحية والثقافية والرياضية والدينية (العامة والخاصة) • قسم الشباب الراضي عن مستوى الخدمات • أسباب الرضا وعدم الرضا

نموذج مؤشرات لتقييم الشباب

البيئة المحلية		
الفئة	المؤشرات المختارة	
الاقتصاد المحلي	• تقسيم الشركات القائمة حسب القطاع/الصناعة • متوسط الأجر الشهري حسب الصناعة (بالدولار) • الطلب على العمالة حسب الصناعة • تصورات أرباب العمل حول الشباب • فجوات المهارات المذكورة من قبل أرباب العمل	
الحكومة المحلية	• مخصصات ميزانية حسب القطاع (برامج البنية التحتية والتعليم والعدالة والبرامج الخاصة بالشباب وغيرها) • المبادرات العامة الحالية التي تستهدف الشباب	
الحصول على التكنولوجيا ووسائل الإعلام واستخدامها	• نسبة الشباب الذين يستخدمون الهواتف المحمولة • نسبة الشباب الذين تتوافر لديهم حواسيب واتصال بشبكة الانترنت • نقان المستخدم للإزاعة والتلفاز وغيرها • كيف يتم استخدام التكنولوجيا ووسائل الإعلام؟ وفيم يتم استخدامهم؟	
الحصول على الخدمات المالية واستخدامها	• عدد ونسبة الشباب الذين لديهم حساب بنكي • أنواع حسابات التوفير المتاحة • مصادر الائتمان • شروط اقتراض الأموال من البنوك المحلية (الحد الأدنى للسن والضمانات وغيرها)	
البيئة الاجتماعية		
الفئة	المؤشرات المختارة	
ديموغرافيا البلد	• عدد السكان حسب العمر والجنس • النسبة المئوية للسكان الذين تقل أعمارهم عن ٢٥/٢٩ • النسبة المئوية للسكان الذين تتراوح أعمارهم بين ١٢/١٥ و٢٤/٢٩ • نسبة الذكور إلى الإناث • النسبة المئوية لسكان المناطق الحضرية والريفية • النسبة المئوية للسكان الشباب حسب الدين والعرق • النسبة المئوية للسكان الشباب حسب اللغة الأساسية	
الإطار القانوني	• سن الرشد • سن التصويت • الحد الأدنى للسن القانونية للزواج • الحماية القانونية والالتزامات حسب السن • القانون العرفي	
المعتقدات والفكر المهيمن	• التنظيم والتسلسل الاجتماعي الهرمي • نسب العائلة • أنماط الزواج • أنماط الميراث • نظام القيم	

ملاحظة: للحصول على مزيد من المعلومات المتباينة، يمكن تفصيل العديد من المؤشرات المذكورة أعلاه حسب العمر والجنس ومستوى التعليم ودخل الأسرة المعيشية وغيرها.

ملحق ٢: حلول التكاليف

كثيراً ما يفيد القليل من الابتكار في تقليص تكاليف دراسة الأثر. إلا أن ذلك ينطوي أيضاً على مقايضة بجودة التقييم، لذا ينبغي مراعاة كل النواحي باهتمام.

تقليص تكاليف الخبراء الاستشاريين. ليس من الضروري دائماً الاستعانة بخبير دولي لقيادة التقييم. فالعمل مع أخصائي من خلال منظمة أو جامعة شريكة (حيث يمكن للخبير الاستشاري استخدام نتائج التقييم المنشور بحثي) قد يوفر الكثير من الأموال. إضافة إلى ذلك، كثيراً ما يكون طلاب الماجستير والدكتوراه في مجالات الاقتصاد والصحة العامة أو غيرها من العلوم الاجتماعية مزودين بالخلفية التحليلية اللازمة لدعم التقييم وتصميم المسوح وإجراء تحليل البيانات وسيسعدهم العمل في برامج من الواقع.

تقليص تكلفة جمع البيانات. يتم تحديد إمكانية ومدى تقليص تكلفة جمع البيانات بالتعاون مع خبير التقييم الرئيسي عند إتمام خطة ومنهجية التقييم.

- **الاعتماد على البيانات القائمة.** قد لا يكون من الضروري أيضاً جمع بياناتنا إذا كان بالإمكان التأسيس على المسوح المحلية أو الإقليمية أو الوطنية التي قد تكون أجريت في الماضي أو تلك التي ما زالت في مرحلة التخطيط، إذا كانت هناك بيانات مسوح متاحة، يمكن أن يقتصر تقييم الأثر على تحليل البيانات. على العكس، إذا كان هناك مسح تم التخطيط له بالفعل وسيغطي مجتمع الدراسة موضع اهتمامنا، قد يتمكن التقييم من إضافة سلسلة من الأسئلة ذات الصلة ببرنامجنا.
- **تقليص حجم العينة.** تجرب بعض البرامج تقليل عدد الأفراد الخاضعين للمسح من أجل توفير الأموال. بالرغم من أن هذا يقلص التكاليف، إلا أنه أيضاً يزيد من احتمالية ألا يجد التقييم عن طريق الخطأ أي أثر للبرنامج مع أنه يوجد في الواقع. يجب استخدام عينات صغيرة من الأفراد بحذر. سوف يكون خبير التقييم المعين للإشراف على تقييم الأثر قادراً على التوصية بالحجم الأدنى للعينة التي يمكن استخدامها في السياق المحدد.

- **تقليص طول أداة المسح.** تميل استبيانات المسح إلى أن تكون طويلة حيث يقترح مختلف أصحاب المصالح بنود إضافية قد يوجد اهتمام بإضافتها. بحصر الاستبيان في النتائج الرئيسية موضع الاهتمام وترك الأسئلة التي لا ترتبط ارتباطاً مباشراً بالأهداف الجوهرية للتدخل، يمكن توفير الوقت والمال أثناء جمع وتحليل البيانات.

- **استخدام أدوات قائمة.** ليس من الضروري دائماً وضع مسح جديد تماماً لكل تقييم. إذا كان لدى منظمة تصميم برنامج موحد عبر كل التدخلات، فمن الممكن استخدام استبيانات قائمة من داخل أو خارج المنظمة (أو أجزاء منها على الأقل) والتي قد تم التحقق من صدقها بالفعل. كثيراً ما يكون ممكناً بناء مجموعة من الأسئلة ووحدات المسوح والتي يمكن استخدامها في تقييمات متعددة تشترك في نفس منطق وأهداف التدخل.

- **استخدام عملية التسجيل بالبرنامج لجمع بيانات أولية.** في بعض الأحيان يمكن استخدام عملية الطبيعية للتسجيل بالبرنامج لجمع بيانات عن المشتركين. قد تتضمن هذه المعلومات خصائص اجتماعية سكانية أساسية ومعلومات حول نتائج ذات أهمية، والتي يمكن بدورها أن تستخدم كبيانات أولية.

- **جمع البيانات مع العاملين بالبرنامج أو الموظفين المحليين.** يمكن أن يؤدي جمع البيانات داخلياً بدلاً من الاستعانة بشركة مسوح إلى توفير الكثير من المال. وينطبق نفس الأمر على استخدام مدرسين (عندما يكون البرنامج معتمداً على مدرسة أو مركز تدريب)، أو طلاب الجامعات أو غيرهم من المجتمع المحلي المستعدين للعمل بأجر أقل كثيراً مما تتقاضاه شركة إجراء مسوح. قد ينجح هذا الأمر في المسوح القصيرة والبسيطة ولكن له بعض العيوب الهامة، خاصة في حالات جمع البيانات الأعمق والأكبر. نظراً لتعدد جمع البيانات، عادة ما لا يتمتع العاملين بالبرنامج أو الموظفين المحليين بالخبرة والمهارات اللازمة. وقد ترتفع تكاليف الإشراف والتدريب إلى جانب زيادة الوقت اللازم لإكمال جمع البيانات. إذا حدث أي خطأ في مرحلة جمع البيانات، فقد يعني هذا أن البيانات عديمة الجدوى، ويجب أن تعاد عملية جمع البيانات. قد يؤدي ذلك إلى أخطاء باهظة التكلفة. علاوة على ذلك، قد يثير جمع البيانات مع العاملين بالبرنامج مخاوف حول الحيادية وبالتالي الثقة في النتائج.
- **استخدام أدوات مبتكرة لجمع البيانات.** مع استمرار ظهور التقنيات الحديثة، يتزايد عدد التقنيات الحديثة التي قد تستخدم في جمع البيانات. على سبيل المثال، بدلاً من استخدام مسوح ورقية، يمكن اللجوء إلى إدخال البيانات باستخدام الهواتف المحمولة أو الحاسوب على الفور، وبالتالي تقليل الوقت والتكلفة اللازمين لاحقاً لعمليات إدخال البيانات ومعالجتها حرفياً إلى الصفر.
- **استخدام الاستبيانات الذاتية أو المقابلات الهاتفية.** بالرغم من أن ذلك يقلل الكثير من التكاليف، إلا أن الاعتماد على الاستبيانات الذاتية بدلاً من الاستبيانات التي يديرها مجري المقابلات لا يعتبر عادة إستراتيجية عملية لجمع البيانات لوضع برامج سبل عيش الشباب في الدول النامية، وخاصة مع انخفاض المستوى التعليمي. لذلك يجب استخدامها في ظروف محددة فقط وبعد إجراء مشاورات مكثفة مع خبير التقييم.
- **التركيز على المناطق الجغرافية التي يسهل الوصول إليها.** يمكن توفير التكاليف بقصر جمع البيانات على مناطق جغرافية يسهل الوصول إليها. إذا كان أحد البرامج ينفذ في مواقع مختلفة في أنحاء البلد، يمكن أن يركز التقييم على العاصمة بدلاً من المناطق الريفية النائية. على الرغم من ذلك، يعني هذا أن النتائج ستمثل فقط فئة سكانية فرعية معينة ومنطقة جغرافية معينة (مثل منطقة حضرية) يشملها التقييم.
- **تقليص المسح الأولي.** ليس من الضروري دائماً إجراء مسح أولي كبير أو، في بعض الحالات، أي مسح أولي على الإطلاق. على سبيل المثال، عندما يكون البرنامج معد جيداً وله إطار عمل متابعة قوي، يكفي إجراء مسح أولي صغير يجمع بيانات سريعة حول المشاركين لفهم وضعهم الحالي ومقارنته عبر المجموعتين. بالمثل، بالنسبة للتقييمات العشوائية، إذا استخدم أحد التقييمات عينة كبيرة جداً، لن يكون هناك داع للقلق بشأن الاختلافات بين مجموعة المعالجة ومجموعة المقارنة، وبالتالي ليس هناك ضرورة على الإطلاق لمسح أولي.

بالرغم من ذلك، من الخطورة إهمال المسح الأولي بالكامل. فالمسح الأولي يعني اتصال أولي مع المستجيبين، مما يساعد كثيراً في التأكيد على الفئة السكانية التي يتعين على المنسقين متابعتها. كما أنه يتضمن أيضاً دقة معلومات المواقع من أجل العثور على الأفراد لغرض مسح المتابعة. يميل الشباب إلى الارتحال كثيراً، لذلك، فعدم إجراء مسح أولي قد يزيد كثيراً من تكلفة العثور على هؤلاء الأفراد خلال عام أو عامين أو ثلاثة أعوام. بدون مسح أولي، يستحيل أيضاً استخدام القائمة الكاملة للخصائص الفردية المستخدمة بشكل نموذجي لتحليل البيانات. يمكن أن يظل استخدام النوع، إلا أن معرفة نوعية حياة الأفراد في البداية لن يكون ممكناً.

طلب تمويل خارجي. ليس من اللازم أن يكون تمويل تقييم الأثر مغطى بالكامل من ميزانية البرنامج. في الواقع، كثيراً ما يكمل التمويل بمصادر خارجية، مثل الحكومة أو الأمم المتحدة أو البنوك التنموية أو المؤسسات أو المنظمات الخيرية أو منظمات البحوث والتقييم. في الواقع، أدت الزيادة المتنامية في وضع البرامج والسياسات القائمة على أدلة إلى توافر التمويل من مجموعة متنوعة من المصادر. عندما يكون لدى التقييم القدرة على سد ثغرة بحثية كبيرة تهتم المجتمع التنموي بأكمله، كما في حالة معظم تقييمات الأثر في مجال سبل عيش الشباب، يكون الممارسون على ثقة بأنه بالإمكان توفير تمويل خارجي.

تأجيل جمع بيانات المتابعة. في بعض الحالات، لا تكفي كل مصادر التمويل الداخلية والخارجية معاً لتمويل كل التقييم الذي قد يشمل دورات عديدة محتملة لجمع البيانات. في هذه الحالة، قد يقرر فريق البرنامج والمقيم إقتصار الإلتزام الأولي على تصميم تقييم صامد والمسح والتحليل الأوليين. عند توافر التمويل في المستقبل، يمكن عندئذ إضافة جمع وتحليل بيانات المتابعة. في الواقع العملي، كثيراً ما يكون الحصول على تمويل إضافي لتمويل جمع بيانات المتابعة بالمقارنة أسهل من محاولة الحصول على تمويل خارجي من أجل التقييم برمته.

ملحق ٣: اختبارات التحقق والتزيف

اليناصيب العشوائي والتنفيذ على مراحل

يعتبر التخصيص العشوائي من أكثر الأساليب الصامدة لتقدير الواقع المضاد، وهو بمثابة المعيار الذهبي لتقييم الأثر. مع ذلك، ينبغي مراعاة استخدام بعض الاختبارات الأساسية لتقييم صدق استراتيجية التقييم هذه في سياق معين.

- هل الخصائص الأولية متوازنة؟ قارن الخصائص الأولية لمجموعة المعالجة ومجموعة المقارنة.
- هل حدث أي عدم امتثال للمهمة؟ تأكد مما إذا كانت كل الوحدات المستحقة قد حصلت على المعالجة وأنه ما من وحدات غير مؤهلة قد حصلت على المعالجة. إذا تبين وجود عدم امتثال، استخدم طريقة العرض العشوائي.
- هل أعداد الوحدات في مجموعات المعالجة والمقارنة كبيرة بما يكفي؟ إذا لم تكن كذلك، قد نرغب في الجمع بين طريقة التخصيص العشوائي والاختلاف في الاختلاف.

الترويج العشوائي

يؤدي الترويج العشوائي إلى تقديرات صادقة للواقع المضاد إذا عززت الحملة الترويجية الإنضمام للبرنامج دون التأثير بشكل مباشر على النتائج الذي يعكس الاهتمام.

- هل الخصائص الأولية متوازنة بين هؤلاء الذين تلقوا على الحملة الترويجية والآخرين الذين لم يتلقوها؟ قارن الخصائص الأولية بين المجموعتين.
- هل الحملة الترويجية تؤثر بشكل كبير على الإنضمام للبرنامج؟ يجب أن تؤثر. للتأكد، قارن معدلات الإنضمام للبرنامج في العينة التي تم الترويج فيه والعينة التي لم يتم فيها الترويج.
- هل تؤثر الحملة الترويجية تأثيراً مباشراً على النتائج؟ لا يجب أن تؤثر. لا يمكن اختبار ذلك عادة بشكل مباشر، وبالتالي يجب أن نعتمد على النظرية والمنطق لتوجيهنا.

تصميم عدم الاستمرارية

يشترط تصميم الانحدار غير المستمر (RDD) أن يكون مؤشر الأهلية مستمر حول مستوى الحد وأن تكون الوحدات قابلة للمقارنة على مقربة من درجة مستوى الحد.

- هل المؤشر مستمر حول مستوى الحد في المرحلة الأولية؟
- هل ظهر أي عدم امتثال عند مستوى الحد؟ اختبر ما إذا كانت كل الوحدات المؤهلة قد حصلت على المعالجة وأنه ما من وحدات غير مؤهلة قد حصلت على المعالجة. إذا وجدنا عدم امتثال، فيجب أن نجمع الانحدار غير المستمر مع أساليب متقدمة أكثر لتصحيح «عدم الاستمرارية الغامضة».

الاختلاف في الاختلاف

يفترض أسلوب الاختلاف في الاختلاف تشابه اتجاهات النتائج في مجموعتي المقارنة والمعالجة قبل التدخل وأن العوامل الوحيدة التي تفسر التغيرات في النتائج بين المجموعتين تكون ثابتة بمرور الوقت.

- هل يمكن أن تكون النتائج قد تحركت بالتزامن في مجموعتي المعالجة والمقارنة في غياب البرنامج؟ يمكن تقييم ذلك باستخدام اختبارات تزييف متعددة مثل ما يلي: (١) هل تحرك النتائج في مجموعتي المعالجة والمقارنة بالتزامن قبل التدخل؟ إذا أتاحت دورتان لجمع البيانات قبل بدء البرنامج، اختبر ما إذا كان يوجد فرق في الاتجاهات بين المجموعتين. (٢) ماذا عن النتائج المزيفة التي لا يجب أن تتأثر بالبرنامج؟ هل هي تتحرك بالتزامن قبل وبعد بدء التدخل في مجموعتي المعالجة والمقارنة؟
- قم بإجراء تحليل الاختلاف في الاختلاف باستخدام عدة مجموعات مقارنة معقولة. فهل نحصل على تقديرات مشابهة لأثر البرنامج؟

- قم بإجراء تحليل الاختلاف في الاختلاف مستخدماً مجموعات المعالجة والمقارنة المختارة ونتيجة زائفة لا يجب أن تتأثر بالبرنامج. يجب أن يكون تأثير البرنامج على هذا النتاج صفر.
- قم بإجراء تحليل الاختلاف في الاختلاف مستخدماً متغير النتاج المختار على مجموعتين نعرف أنهما لم تتأثرا بالبرنامج. يجب أن يكون تأثير البرنامج صفر.

المطابقة

- تعمد المطابقة على الافتراض أن الوحدات الملتحقة وغير الملتحقة متشابهة من حيث المتغيرات غير الملحوظة التي قد تؤثر على احتمالية المشاركة في البرنامج والنتائج.
 - هل المشاركة في البرنامج تتحدد بمتغيرات لا يمكن ملاحظتها؟ لا يمكن اختبار ذلك بشكل مباشر، لذلك يجب أن نعتمد على النظرية والمنطق.
 - هل الخصائص الملحوظة متوازنة بين المجموعات الفرعية المطابقة؟ قارن الخصائص الملحوظة لكل مجموعة وحدات معالجة ومجموعة وحدات المقارنة المطابقة لها.
 - هل يمكن إيجاد وحدة مقارنة مطابقة لكل وحدة من وحدات المعالجة؟ تحقق من وجود ما يكفي من الدعم العام في توزيع نقاط الميل؟ تشير نطاقات الدعم العام الصغيرة إلى اختلاف كبير بين الأفراد الملتحقين وغير الملتحقين مما يلقي ظلالاً من الشك على إمكانية الوثوق في أسلوب المطابقة.
- المصدر: أعيد نشره بتصريح من غيرتلر وآخرين (٢٠١١، ص. ١١٨-١١٩).

GLOBAL PARTNERSHIP FOR YOUTH EMPLOYMENT

In 2008, with support from the World Bank Development Grant Facility, the International Youth Foundation, the Youth Employment Network, the Arab Urban Development Institute, and the Understanding Children's Work Project joined together to form the Global Partnership for Youth Employment (GPYE). Its goal: to build and disseminate evidence on youth employment outcomes and effective programs to help address the challenges facing young people in their transition to work. The GPYE leverages the technical and regional experience of the five partner organizations in youth employment research, programming, evaluation, and policy dialogue. The partnership's work focuses on Africa and the Middle East, regions in need of better evidence on effective approaches to promote youth employment.

For more information, visit www.gpye.org