

مقدمة إلى تقييم الأثر

تأليف باتريشيا جيه. روجرز
جامعة RMIT (أستراليا) و BetterEvaluation



تتعلق هذه الملاحظة الإرشادية الأولى من سلسلة ملاحظات من أربعة أجزاء بتقييم الأثر وقد تم تطويرها من قبل مؤسسة InterAction مع دعم مالي من مؤسسة Rockefeller Foundation. الملاحظات الأخرى في هذه السلسلة هي: رصد الربط وتقييم تقييم الأثر، ومقدمة إلى الطرق المختلطة في تقييم الأثر، واستخدام نتائج تقييم الأثر. يمكن الاطلاع على السلسلة الكاملة في موقع InterAction على الرابط: <http://www.interaction.org/impact-evaluation-notes>.

الشكر والتقدير

استفادت هذه الملاحظة الإرشادية من الإفادة الامرتجاعية لأعضاء مجموعة InterAction's Evaluation and Program Effectiveness Working Group. واستُمدت المعلومات حول طرق التقييم في الملاحظة الإرشادية 1 من BetterEvaluation، وهو عبارة عن تعاون دولي لمشاركة المعلومات من أجل تحسين التقييم. يمكن الاطلاع على المزيد من المعلومات حول كل من هذه الطرق على الموقع الإلكتروني www.betterevaluation.org.

1	مقدمة
2	1. ماذا نعني بـ "تقييم الأثر"؟
2	2. لماذا يجب أن نقوم بإجراء تقييم للأثر؟
3	3. ما هي الأسئلة التي يحاول تقييم الأثر الإجابة عنها؟
3	4. ما هي الجهة التي يجب أن تقوم بإجراء تقييم الأثر؟
3	5. كيف ينبغي أن نختار طرق تقييم الأثر؟
4	6. توضيح القيم وراء تقييم الأثر
5	7. تطوير نظرية أو نموذج للكيفية المفترضة لعمل التدخل
7	8. قياس أو وصف الآثار (والتغيرات الأخرى المهمة)
8	9. شرح مدى إنتاج التدخل للنتائج الملحوظة
11	10. صنع الأدلة
11	11. رفع التقارير بالنتائج ودعم استخدامها
12	12. متى يجب إجراء تقييم الأثر؟
12	13. ما هو المطلوب لإجراء تقييم أثر ذي جودة؟
13	14. التحديات الشائعة في تطوير تقييم الأثر
15	الخلاصة
15	المراجع والموارد المفيدة الأخرى

مقدمة

يمكن أن يحسّن تقييم الأثر الملائم والموثوق به بشكل كبير من فعالية التطوير. لقد أدى التأكيد المتزايد على تقييم الأثر في التطوير إلى أسئلة كثيرة. ما الذي يشكل تقييم الأثر الملائم والموثوق به؟ كيف ينبغي أن يُدار تقييم الأثر؟ ما هي الإجراءات ومصادر البيانات المناسبة؟ كيف يمكن الجمع بين البيانات الكمية والنوعية بشكل فعّال في تقييم الأثر؟ ما الذي يجب عمله لدعم الاستخدام الملائم لتقييمات الأثر؟ ما هي المعاني الضمنية للتركيز المتزايد على تقييم الأثر للأنواع الأخرى من الرصد والتقييم؟

وتصف الطرق والأساليب والتصاميم المختلفة التي يمكن استخدامها للجوانب المختلفة لتقييم الأثر: توضيح قيم التقييم، وتطوير نظرية حول كيفية فهم أداء التدخل، وقياس أو وصف الآثار والمتغيرات المهمة الأخرى، وشرح سبب حدوث الآثار، وصنع النتائج، ورفع التقارير ودعم الاستخدام. وتناقش الملاحظة ما يعتبر تقييم أثر جيد - وهو التقييم الذي يحقق التوازن بين الضرورات المتنافسة في أن يكون مفيداً ودقيقاً وأخلاقياً وعملياً - وكيفية تحقيق ذلك. تحتوي الحواشي السفلية في كامل الوثيقة على مراجع للقراءة الإضافية في مجالات معينة.

لقد أنتجت InterAction سلسلة من الملاحظات الإرشادية تتناول هذه الأسئلة وذلك لدعم كادر الإدارة والبرنامج والمراقبة والتقييم في المنظمات الدولية غير الحكومية لتخطيط تقييمات الأثر وإدارتها وإجرائها واستخدامها. ويمكن لهذه الملاحظات أن تشكل مناقشاتهم مع المقيّمين والشركاء والممولين الخارجيين.

تقدم الملاحظة الإرشادية الأولى، "مقدمة إلى تقييم الأثر"، نظرة عامة حول تقييم الأثر وتشرح كيف أن تقييم الأثر يختلف عن الأنواع الأخرى للتقييم ويتممها وسبب وجوب إجراء تقييم الأثر وموعده والجهة التي سوف تقوم به.

1. ماذا نعني بـ "تقييم الأثر"؟

يتحرى تقييم الأثر التغيرات التي يحدثها التدخل. ويمكن أن يُجرى تقييم الأثر على التدخلات من أي حجم: مشروع محلي صغير لفيروس نقص المناعة البشرية-الايديز، أو برنامج تعزيز للمجتمع المدني بأكمله خاص بمنظمة غير حكومية، أو سلسلة من مشاريع إدارة الموارد الطبيعية تُجرى في منطقة جغرافية، أو مجموعة من الأنشطة المتزامنة التي تقوم بها منظمات مختلفة والتي تهدف إلى تحسين قدرات المجتمع.

تعتبر النتائج المتوقعة من التدخل جزءاً مهماً من تقييم الأثر، ولكن من المهم أيضاً أن يتم فحص النتائج غير المتوقعة. تُعرّف التدخلات في هذه الملاحظة الإرشادية بالتالي:

النتائج الإيجابية والسلبية، والمقصودة وغير المقصودة، والمباشرة وغير المباشرة، والرئيسية والثانوية التي يُحدثها التدخل. تعريف لجنة المساعدة الإنمائية بمنظمة التعاون والتنمية الاقتصادية (OECD)¹

تُفهم الآثار عادة بأنها تنشأ بعد النتائج الوسيطة ونتيجة لها. على سبيل المثال، تحقيق النتائج الوسيطة للوصول المحسّن للأرض والمستويات المتزايدة للمشاركة في صنع القرار في المجتمع قد يحدث قبل، ويسهم في، الأثر الأخير المقصود وهو الرفاهية والصحة المحسّنة للمرأة. يمكن أن يكون الفارق بين النتائج والآثار نسبياً ويعتمد على الأهداف المعلنة للتدخل.

وعملياً، غالباً ما يكون من المفيد أن يشمل التقييم كلاً من النتائج والآثار. فذلك يعطي مؤشرات مبكرة لمعرفة نجاح التدخل من عدمه - وإذا لم يكن التدخل ناجحاً فسيساعد ذلك في تحديد مكان الخطأ وربما سببه.

في هذه الملاحظة الإرشادية، يشمل تقييم الأثر أي تقييم يفحص بشكل منهجي وتجريبي الآثار التي أحدثها التدخل. يستخدم بعض الأفراد والمنظمات تعريفاً أضيق لتقييم الأثر، ويشملون فقط التقييمات التي تحتوي على واقع مضاد من نوع ما (تقديراً لماذا كان سيحصل لو أن التدخل لم يحدث) أو نوعاً معيناً من الواقع المضاد (مثلاً، مقارنات بمجموعة لم تتلق التدخل). فعلى سبيل المثال، تستخدم الوكالة الأمريكية للتنمية الدولية التعريف التالي: "تقيس تقييمات الأثر التغير في نتيجة تطور يمكن نسبتها إلى تدخل محدد. تركز تقييمات الأثر على نماذج للسبب والنتيجة وتتطلب واقعاً مضاداً موثقاً به ومعرفاً بدقة لضبط العوامل بخلاف التدخل الذي قد يفسر التغير الملحوظ". هذه التعريفات المختلفة مهمة حين نقرر ماهية الطرق أو تصاميم البحث التي ستعتبر موثوقة من قبل مستخدمي التقييم المقصودين أو من قبل الشركاء أو الممولين.

وتقييم الأثر بالطبع ليس النوع الوحيد للتقييم الذي يدعم التطوير الفعّال. من المهم أن نضمن أن الاستثمارات في تقييم الأثر (من ناحية الوقت والمال) لا تتم على حساب الرصد أو الأنواع الأخرى للتقييم - مثل تقدير الحاجات وتقييم العملية وتقييم التكلفة-المنفعة - التي تعد ضرورية أيضاً لتشكيل القرارات حول الممارسة والسياسة. تناقش الملاحظة 2 كيفية إمكان إجراء تقييم الأثر وتلك الأنواع الأخرى من الرصد والتقييم بطرق تدعم بعضها البعض. على سبيل المثال، يمكن أن يقدم رصد البيانات أساساً جيداً لتقييم الأثر، ويمكن لتقييم الأثر أن يوجه تطوير أنظمة الرصد. ويقدم تقييم الأثر المعلومات الضرورية لتقييم التكلفة-المنفعة وتقييم التكلفة-الفعالية.

2. لماذا يجب أن نقوم بإجراء تقييم للأثر؟

إن أفضل طريقة للبدء في تقييم معين للأثر تعتمد جزئياً على الهدف من التقييم وعلى المستخدمين المقصودين الرئيسيين. تتضمن بعض الأسباب الشائعة للقيام بعمل تقييم الأثر الأمور التالية:

- اتخاذ القرار حول ما إذا سيمول التدخل أم لا - يتم إجراء "التقييم السابق" قبل تنفيذ التدخل لتقدير آثاره المحتملة وتشكيل القرارات المتعلقة بالتمويل.

¹ تُعرّف الآثار أحياناً بطريقة مختلفة تماماً. على سبيل المثال، يميز دليل تطوير النموذج المنطقي لمؤسسة ديليو، كيه. كيلوغ (www.wkkf.org/knowledge-center/resources/2006/02/WK-Kellogg-Foundation-Logic-Model-Development-Guide.aspx)

الأثر من ناحية انتشاره إلى ما وراء أولئك المشاركين مباشرة في البرنامج. ويُشار إلى التغييرات الجذدة في سلوك المشاركين في البرنامج ومعرفةهم ومهاراتهم وحالاتهم ومستوى أدائهم على أنها «النتائج». وتوصف فقط التغييرات في المؤسسات أو المجتمعات أو الأنظمة كنتيجة لأنشطة البرنامج خلال سبعة إلى 10 سنوات بأنها «آثار».

ولكن، بالنسبة لبعض أصحاب المصالح، لا يُنظر إلى المقيمين الخارجيين دائماً بأنهم غير متحيزين بما أن جمعهم وتفسيرهم للبيانات قد يتأثر بافتقارهم للاعتياد على السياق. وفي بعض الحالات، دفع أصحاب المصالح من البرنامج و/أو أعضاء المجتمع إلى المشاركة في إجراء التقييم يمكن أن يضيف الدقة والمصداقية عن طريق دعم وصول أفضل للبيانات (خصوصاً المعلومات الأساسية) وتفسير أكثر ملاءمة للبيانات.

ويمكن لثلاث ممارسات على الأخص أن تنتج في الغالب تقيماً هو الأفضل من حيث الجودة: إنشاء فريق من المقيمين ذوي رؤى خارجية وداخلية، وضمان الشفافية من ناحية البيانات التي يجري استخدامها وكيفية استخدامها في التقييم، والتثليث - أي استخدام مصادر متعددة للأدلة (والتي تحمل نقاط قوى مكتملة) ورؤى متعددة في التحليل والتفسير. ومن المفيد على وجه الخصوص أن نلحق بالفريق خبراء التقييم المحليين الذين يعرفون السياق والتاريخ والتدخلات المقارنة من قبل الوكالات الأخرى.

من الممكن إدارة التقييم من قبل مجموعة داخلية (ربما تكون لجنة توجيه داخلية تُشكل من قبل مجموعة استشارية ذات عضوية خارجية) أو من قبل مجموعة موحدة. تستخدم الأساليب التشاركية لإدارة التقييمات عادة طاقم البرنامج وأعضاء المجتمع وشركاء التنمية. فهم لا يشاركون في جمع البيانات فحسب، بل يشاركون أيضاً في التفاوض على الغرض من تقييم الأثر وتطوير أسئلة التقييم الرئيسية وتصميم تقييم يجب عليها ومواصلة العمل على النتائج.²

5. كيف ينبغي أن نختار طرق تقييم الأثر؟

لقد كان هناك نقاش كبير حول تقييم التطوير، وبشكل أوسع، حول أفضل الطرق لتقييم الأثر. تعكس هذه

- اتخاذ القرار حول ما إذا سيتم الاستمرار في التدخل أو توسيعه أم لا.
- تعلم كيفية تكرار أو توسيع تدخل تجريبي.
- تعلم كيفية تهيئة تدخل ناجح ليناسب سياقاً آخر بنجاح.
- طمأنة الممولين، بمن فيهم المتبرعين ودافعي الضرائب (المساءلة التصاعديّة)، بأن المال يُستثمر بحكمة - ويشمل ذلك تعلم المنظمة للأمور التي تنجح والأمور التي لا تنجح واستخدامها لهذه المعلومات لتحسين التنفيذ المستقبلي وقرارات الاستثمار.
- تبليغ المستفيدين والمجتمعات المقصودة (المساءلة التنافسية) حول ما إذا كان البرنامج ينفذ المجتمع أم لا وبأي الطرق ينفذه.

تناقش الملاحظة الإرشادية 4 بتفصيل أكبر كيفية دعم هذه الطرق المختلفة لاستخدام تقييم الأثر.

3. ما هي الأسئلة التي يحاول تقييم الأثر الإجابة عنها؟

يجب أن يركز تقييم الأثر على عدد صغير من أسئلة التقييم الرئيسية (خمسة إلى سبعة أسئلة). وهذه هي الأسئلة ذات المستوى العالي التي يتناولها التقييم، وليست الأسئلة المحددة التي قد يتم طرحها في مقابلة أو استبيان. فمن الأفضل التركيز على عدد صغير من الأسئلة التي تتعلق بشكل مباشر بالهدف بدلاً من نشر موارد التقييم وتركيز المستخدمين على عدد كبير من الأسئلة.

4. ما هي الجهة التي يجب أن تقوم بإجراء تقييم الأثر؟

يمكن أن تتولى الجهات التالية إجراء تقييم الأثر: مُقيّم أو فريق تقييم خارجي، أو وحدة من المنظمة القائمة على التنفيذ تكون داخلية ولكن منفصلة، أو المشاركون في التدخل (من فيهم أعضاء المجتمع)، أو فريق موحد من المقيمين الداخليين والخارجيين.

يمكن أن يقدم المقيّم الخارجي مجموعة من الخبرات والتجارب التي قد لا تكون متوفرة داخل المنظمة ويكون لديه استقلالية ومصداقية أكبر من المقيّم الداخلي. على سبيل المثال، تضع سياسة التقييم الخاصة بالوكالة الأمريكية للتنمية الدولية توقعاً بأن معظم التقييمات سيقوم بها مقيّمون خارجيون.

² تشمل المصادر الإضافية على الطرق التشاركية ما يلي: Marisol Estrella et al. (eds), Learning from Change: Issues and Experiences in Participatory Monitoring and Evaluation (Brighton: Institute of Development Studies, 2000), <http://www.idrc.ca/EN/Resources/Publications/Pages/IDRCBookDetails.aspx?PublicationID=348>; Andrew Catley et al., "Participatory Impact Assessment," Feinstein International Center, Tufts University: October 2008, <http://sites.tufts.edu/feinstein/2008/participatory-impact-assessment>; Robert Chambers, "Who Counts? The Quiet Revolution of Participation and Numbers," Working Paper No. 296 (December 2007), Brighton: Institute of Development Studies, <http://www.ids.ac.uk/files/Wp296.pdf>.

أمثلة على أسئلة التقييم الرئيسية لتقييم الأثر

- ما الذي ساعد أو أعاق [التدخل] عن تحقيق هذه الآثار؟
- الأثر الإجمالي
- هل نجح الأمر؟ هل قام [التدخل] بإحداث [الآثار المقصودة] في المدى القصير والمتوسط والبعيد؟
- لمن وبأي الطرق وفي أية ظروف نجح [التدخل]؟
- ما هي الآثار غير المقصودة (الإيجابية والسلبية) التي أحدثها [التدخل]؟
- طبيعة الآثار وتوزيعها
- هل يُرجح أن تكون الآثار مستدامة؟
- هل وصلت هذه الآثار إلى كل المستفيدين المقصودين؟
- تأثير العوامل الأخرى على الآثار
- كيف عمل [التدخل] بالاشتراك مع التدخلات أو البرامج أو الخدمات الأخرى لتحقيق النتائج؟
- كيف ينجح الأمر
- كيف ساهم [التدخل] في [الآثار المرجوة]؟
- ما هي المميزات الخاصة [للتدخل] التي أحدثت فرقاً؟
- ما هي المتغيرات التي كانت موجودة أثناء التنفيذ؟
- كيف كانت جودة التنفيذ في المواقع المختلفة؟
- إلى أي مدى تم تفسير الفروق في الأثر بواسطة المتغيرات في التنفيذ؟
- مطابقة الآثار المقصودة للحاجات
- إلى أي مدى تطابقت الآثار مع حاجات المستفيدين المقصودين؟

- التغيير أو نماذج المنطق أو نظرية البرنامج.
- قياس أو وصف هذه الآثار والمتغيرات الأخرى ذات الصلة، بما في ذلك العمليات والسياق.
- شرح ما إذا كان التدخل سبباً للآثار الملحوظة.
- صنع الأدلة في حكم تقييمي كلي.
- رفع التقارير بالنتائج ودعم استخدامها.

تناقش هذه الملاحظة الإرشادية كل هذه الأوجه، وتقدم المعلومات حول مجموعة من الطرق التي يمكن استخدامها لهذه الأوجه. تم تقديم روابط لمصادر إضافية للمعلومات. تناقش الملاحظة الإرشادية 3 كيف يمكن لأسلوب الطريقة المختلطة، وهو مزج البيانات الكمية والكيفية بطرق تكاملية، أن يكون قياساً/وصفاً وشرحاً أيضاً.

6. توضيح القيم وراء تقييم الأثر

تتمثل الخطوة الأولى في توضيح القيم التي ستدعم التقييم. يستخلص تقييم الأثر استنتاجات حول مدى نجاح (أو فشل) التدخل، لذا من المهم توضيح حالة النجاح فيما يتعلق بالأمر التالية:

- تحقيق آثار مرغوبة واجتناب (أو على الأقل

النقاشات وجهات نظر مختلفة حول الأمور التي تشكل الأدلة الدقيقة والمفيدة والموثوق بها، ومن الذي ينبغي أن يشارك في إجراء التقييمات والتحكم فيها.

لقد جادلت بعض المنظمات والمقيمين بأنه ينبغي استخدام الطرق الخاصة أو تصاميم البحث كلما أمكن ذلك - مثل التجارب العشوائية المراقبة أو الطرق المشتركة. وجادل آخرون حول مسألة الملاءمة الظرفية. وهذا يعني اختيار الطرق التي تناسب هدف التقييم ونوع أسئلته التي تُطرح وتوفير الموارد وطبيعة التدخل - وعلى الأخص ما إذا كان التدخل موحد القياس أم مهيئ، وما إذا كانت التدخلات تعمل تقريباً بنفس الطريقة في كل مكان ولكل الأشخاص أم أنها تتأثر كثيراً بالسياق.

عند اختيار الطرق، من المهم تناول كل وجه من الأوجه الستة المختلفة لتقييم الأثر:

- توضيح القيم التي ستدعم التقييم - ما الذي سيعتبر مرغوباً فيه أو غير مرغوب فيه من العمليات والآثار وتوزيع التكلفة والمنافع؟
- تطوير و/أو اختبار النظرية حول كيف يفترض أن يعمل التدخل - ويشار إليها أحياناً بنظريات

المختلفة على أنها النتائج أو الآثار الأكثر أهمية.⁴

الطرق التي تساعد في التفاوض بين مجموعات القيم المختلفة

دلفي - وهي عملية تعمل من خلال سلسلة من التفاعلات المكتوبة دون سياق التفاعل وجهاً لوجه حيث يقدم أصحاب المصالح الرئيسيون آراءهم حول ما يرون أنه مهم ثم يستجيبون للنتائج الكلية.⁵

التصويت بالنقاط اللاصقة - في الاجتماعات التي تجري وجهاً لوجه، يقوم الأفراد بتقسيم "أصواتهم" المتعددة (في صورة نقاط لاصقة) على الاختيارات.⁶

7. تطوير نظرية أو نموذج للكيفية المفترضة لعمل التدخل

من المفيد غالباً أن يتم إسناد تقييم الأثر إلى نظرية أو نموذج لكيفية فهم إحداث التدخل لآثاره المقصودة. قد يسمى ذلك نظرية البرنامج أو نظرية التغيير أو سلسلة النتائج أو نموذج المنطق. من الأفضل تطوير نظرية التغيير كجزء من تخطيط التدخل ومن ثم مراجعتها وتعديلها عند الضرورة وعند تخطيط تقييم الأثر. إذا لم ينفذ هذا الأمر في وقت بداية التدخل، فمن الممكن تطوير نظرية تغيير متفق عليها بأثر رجعي.

واعتماداً على موعد تطوير نظرية التغيير، يمكنها أن تعتمد على مجموعة من المصادر: الوثائق الرسمية والأهداف المعلنة، أو البحث في التدخلات المشابهة، أو ملاحظة التدخل أو التدخلات المشابهة، أو سؤال أصحاب المصالح المختلفين (من فيهم القائمين على التخطيط والطاقم والمستفيدين المقصودين) عن رأيهم في عمل التدخل (أو فيما يجب أن يكون عليه عمل التدخل).

4 للمزيد من الإطلاع:

Most Significant Change (MSC), <http://mande.co.uk/special-issues/most-significant-change-msc/>.

5 للمزيد من الإطلاع:

Delphi Method, http://www.unido.org/fileadmin/import/16959_DelphiMethod.pdf; The Delphi Method: Techniques and Applications, <http://is.njit.edu/pubs/delphibook/>; Delphi Survey - Europa, http://ec.europa.eu/regional_policy/sources/docgener/evaluation/evalsed/downloads/sb2_delphi_survey.doc.

6 للمزيد من الإطلاع:

NRCOI Quick Tip, <http://centerforchildwelfare.fmhi.usf.edu/kb/bpam/NRCOI-StickyDotTechnique.pdf>.

تقليل) الآثار السلبية. على سبيل المثال، هل سيتم الحكم على نجاح مشروع تطوير طريق من ناحية الوصول المتزايد إلى الأسواق أو الوصول المحسّن إلى مستشفيات الولادة؟ ما هو مستوى فقدان البيئة الطبيعية والتنوع البيئي الذي سيتم أخذه في الاعتبار كتكلفة معقولة للطريق؟ ما هو المستوى الذي لن يكون بمثابة تنازل مقبول؟

• تحقيق توزيع مرغوب للمنافع. على سبيل المثال، هل يجب علينا أن نحكم على النجاح من ناحية النتيجة التعليمية المتوسطة أو التحسينات المقدمة للمجموعة الأقل حظاً أو رفع المجموعة الضعيفة أو الأقل حظاً (مثل الفتيات الصغار) إلى نفس مستوى نظرائها الأكثر حظاً؟

تعتبر الأهداف المعلنة رسمياً (بما في ذلك الأهداف الإنمائية للألفية) والسياسات التنظيمية بداية مهمة لتوضيح القيم، ولكنها عادة غير كافية في حد ذاتها. فأصحاب المصالح المختلفون قد تكون لديهم في الواقع رؤى مختلفة حول ماهية القيم التي يجب استخدامها في التقييم.

بعض الطرق لتوضيح القيم وراء تقييم الأثر:

الطرق التي تساعد الناس في تبين القيم الضمنية

التقصي التثميني - يستحضر أصحاب المصالح الأساسيون (من فيهم طاقم البرنامج) عندما نجح البرنامج بشكل جيد على وجه الخصوص ثم يحددون القيم التي كان مثلاً عليها خلال تلك الأوقات.³

الدراسات المسحية للمجتمع - يقوم الأفراد في المجتمع إما بتثمين أو تصنيف الأمور التي يرون أنها الأكثر أهمية والتي يجب التطرق إليها.

التغيير الأكثر أهمية - عملية منظمة لتوليد واختيار قصص التغيير التي تحدد الأمور التي يراها الأفراد والمجموعات

3 للمزيد من الإطلاع:

Hallie Preskill, "Using Appreciative Inquiry in Evaluation Practice," <http://www.eval.org/summerinstitute07/handouts/sio7.preskillf.pdf>.

قد تكون هناك نظريات تغيير متعددة - وهي نظريات مختلفة تظهر كيف يعمل التدخل في مراحل مختلفة، وفي سياقات مختلفة (تعترف بمفعول التأثيرات الخارجية) ولآثار مختلفة، ونظريات مختلفة تتطور عبر الزمن بسبب تطور فهمها على نحو أفضل.

يمكن أن تقوم نظريات التغيير على تحسين تقييم الأثر عن طريق المساعدة في:

- تحديد النتائج المتوسطة أو الآثار التي يمكن ملاحظتها خلال الإطار الزمني للتقييم، والتي تكون بدايات للآثار طويلة المدى التي يُقصد من التدخل إحداثها.
- تحديد مكان فشل أو توقف التدخل عن العمل في العملية في حالة عدم نجاح التدخل.
- التمييز بين فشل التنفيذ (حيث لم يتم تحقيق الآثار لأن التدخل لم يُنفذ كما ينبغي) وبين فشل النظرية (حيث لا يؤدي التدخل إلى الأثر المرغوب حتى عندما ينفذ تنفيذاً جيداً).
- تحديد الأوجه التي تجعل التدخل يعمل، وبالتالي تعد هذه الأوجه هامة ومن الضروري الاستمرار فيها عندما يتم تهيئة التدخل لأوضاع أخرى.
- تحديد المتغيرات السلوكية والسياقية المهمة التي ينبغي تناولها في جمع البيانات وتحليلها ورفع التقارير عنها وذلك لفهم المتغيرات في الآثار.
- تقديم إطار عمل مفاهيمي لضم الأدلة المتنوعة حول برنامج يستخدم عدداً كبيراً من التدخلات المتنوعة.

بعض الطرق لتمثيل نظرية التغيير:

أسلوب إطار العمل المنطقي - وهو الشكل الكلاسيكي المستخدم في العديد من منظمات التنمية الذي يستخدم مصفوفة 4x4. الصفوف الأربعة هي الأنشطة والمخرجات والغرض والهدف، والأعمدة الأربعة هي الوصف السردي والمؤشرات القابلة للتحقق منها بشكل

موضوعي ووسائل التحقق والافتراضات.⁷

سلسلة النتائج - يُمثل التدخل بسلسلة من المربعات في تعاقب: المدخلات والأنشطة والمخرجات والنتائج على المدى القصير والنتائج على المدى الأطول والآثار.⁸

سلسلة النتائج /هرمية النتائج/نظرية التغيير - تُمثل النظرية بسلسلة من النتائج المتوسطة التي تؤدي إلى الآثار المقصودة النهائية. يركز هذا الشكل الانتباه على كيفية حدوث التغيير ويفيد في تمثيل البرامج التي تحدث فيها الأنشطة المختلفة بموازاة السلسلة السببية وليس فقط في البداية.⁹

رسم خريطة للنتائج - يركز هذا على تحديد "شركاء الحدود" - وهم المنظمات والمجموعات التي تكون أفعالها وراء نطاق تحكم التدخل، ولكنها ضرورية لتحقيق الأثر - ومن ثم توضيح ما يحتاج هؤلاء الشركاء إلى عمله وكيف يمكن للتدخل أن يسعى للتأثير فيهم.¹⁰

يمكن الاطلاع على موارد مفيدة أخرى لتطوير نظرية التغيير في قسم تطوير نموذج المنطق أو نظرية التغيير.¹¹

يمكن استخدام نظرية التغيير أيضاً في إدارة الآثار السلبية المحتملة أو في تخطيط تقييم الأثر لقياسها.

7 للمزيد من الإطلاع:

The Logical Framework Approach, http://www.ausaid.gov.au/ausguide/pdf/ausguideline3_3.pdf; <http://www.gdrc.org/ngo/logical-fa.pdf>; Logical Framework Analysis, <http://www.gdrc.org/ngo/logical-fa.pdf>; Beyond Lograme: Critique, Variations and Alternatives: <http://www.fasid.or.jp/shuppan/hokokusho/pdf/h213-.pdf>.

8 للمزيد من الإطلاع:

Results Chain: Enhancing Program Performance with Logic Models Guide, <http://www.uwex.edu/ces/lcourse/>; W.K. Kellogg Foundation Logic Model Development Guide, <http://www.evaluationtoolsforracialequity.org/evaluation/tool/doc/pub3669.pdf>.

9 للمزيد من الإطلاع:

Theories of Change and Logic Models: Telling Them Apart, http://www.evaluationtoolsforracialequity.org/evaluation/resource/doc/TOCs_and_Logic_Models_forAEA.ppt.

10 للمزيد من الإطلاع:

Outcome Mapping, <http://publicwebsite.idrc.ca/EN/Programs/Evaluation/Pages/IDRCBookDetails.aspx?PublicationID=121>; Outcome Mapping: ILAC Brief 7, http://www.outcomemapping.ca/download.php?file=/resource/files/csette_en_ILAC_Brief07_mapping.pdf.

11

The Community Toolbox, "Developing a Logic Model or Theory of Change," http://ctb.ku.edu/en/tablecontents/sub_section_main_1877.aspx.

على سبيل المثال، قد يؤدي برنامج معد لتحسين الإنتاجية الزراعية عن طريق تشجيع المزارعين على وضع الأسمدة في حقولهم على زيادة فيضان الفوسفات والضرر البيئي للمجاري المائية. سيقوم تقييم الأثر المتوازن بالتحقيق في هذا الأثر المحتمل بالإضافة إلى الأثر المقصود المتمثل في الإنتاجية المحسنة. يمكن إنشاء نظرية التغيير لفحص كيف يحتمل أن ينتج التدخل آثاراً سلبية، ويمكن استخدام ذلك في تهيئة التدخل من أجل خفض أو تجنب هذه الآثار السلبية، ولتطوير مؤشرات التحذير المبكرة لرصد العمليات، ولضمان إدراج هذه الأمور في خطة تقييم الأثر.

8. قياس أو وصف الآثار (والتغيرات الأخرى المهمة)

يحتاج تقييم الأثر إلى الأدلة الموثوق بها وليس فقط الآثار والمعلومات الجيدة حول مدى حسن تطبيق التدخل مطلوبة أيضاً من أجل التمييز بين فشل التمييز وفشل النظرية. والمعلومات حول السياق مطلوبة أيضاً لفهم إن كان التدخل يعمل فقط في حالات معينة.

من المفيد تحديد أية بيانات متوفرة مسبقاً حول الآثار والتنفيذ والسياسات من المصادر الموجودة، مثل الإحصائيات الرسمية ووثائق البرنامج وأنظمة المعلومات الجغرافية ومشاريع البحث والتقييم السابقة. يمكن جمع المعلومات الإضافية لملء الفجوات أو تحسين جودة البيانات الموجودة باستخدام طرق مثل المقابلات (الفرد والجماعة، المنظمة أو المنظمة جزئياً أو غير المنظمة)، والاستبيانات (بما في ذلك الاستبيانات عبر الإنترنت وجمع البيانات عن طريق الهاتف الجوال)، والملاحظة (المنظمة أو المنظمة جزئياً أو غير المنظمة)، والقياس المباشر (مثلاً لجودة الماء مقارنة بمعيار عالمي).

لا ينبغي أن يذكر وصف الآثار المتوسط فحسب، بل يجب أن يذكر أيضاً مدى تنوع النتائج وبالأخص التبليغ حول الأنماط. يناقش هاورد وايت أهمية النظر إلى التغيرات في مقالته عام 2010:

من المرجح أن تكون الدراسة التي تقدم تقديراً واحداً للآثار (أثر المعالجة المتوسط¹²) أقل فائدة لصناع السياسة من الدراسة التي تفحص أي

12 تأثير العلاج المتوسط هو تقدير الفرق المتوسط الذي يحدثه التدخل. على سبيل المثال، الطلاب في البرنامج بقوا في المدرسة بما متوسطه 2.5 سنة أكثر (مقارنة بمجموعة المراقبة).

السياقات تكون التدخلات فيها أكثر فعالية، والتي تستفيد منها المجموعات الهدف بشكل أكبر، وأي الأوضاع البيئية مفيدة أو ضارة بالنسبة إلى تحقيق الأثر. لذلك يمكن إظهار أن التدخل التعليمي، مثل الألواح الورقية القلابة، يعمل بنجاح، ولكن فقط إذا كان لدى المعلمين أنفسهم مستوى معين من التعليم، أو فقط إذا كانت المدرسة مزودة بشكل جيد مسبقاً بمواد القراءة، أو إذا كان أولياء أمور الطلبة أنفسهم متعلمين.¹³

تشمل بعض المصادر للقياسات والمؤشرات في قطاعات معينة ما يلي:

قائمة أسئلة المسح: شبكة المسح الأسري الدولية - أكثر من 2000 استبيان يمكن البحث فيها حسب الدولة والتاريخ ونوع المسح.¹⁴

الحكم الديمقراطي - مركز الحكم بأوسلو التابع لبرنامج الأمم المتحدة الإنمائي.¹⁵

مؤشر الفقر البشري - ثلاثة مؤشرات تتعلق بالبقاء على قيد الحياة والمعرفة والتموين الاقتصادي التابع لبرنامج الأمم المتحدة الإنمائي.¹⁶

الأهداف الإنمائية للألفية - 48 مؤشراً فنياً و18 هدفاً لـ 8 غايات.¹⁷

التنمية المستدامة - لجنة الأمم المتحدة للتنمية المستدامة 130 مؤشراً للأوجه الاجتماعية والاقتصادية والبيئية والمؤسسية للتنمية

13 Howard White, "A Contribution to Current Debates in Impact Evaluation," Evaluation (April 2010, vol. 16 no. 2): 160.

14 http://www.ihns.org/home/index.php?q=country_questionnaires

15 http://www.undp.org/oslocentre/flagship/democratic_governance_assessments.html

16 <http://hdr.undp.org/en/statistics/indices/hpi/>

17 Millennium Project: <http://www.unmillenniumproject.org/goals/gti.htm>

صحيحة. ما هي الآثار التي حدثت وما هو موعد حدوثها ولمن حدثت؟ هل هذه النتائج متناغمة مع نظرية أن التدخل هو الذي أحدث النتائج أو ساهم في إحداثها؟ المكون الثاني هو الواقع المضاد - أي تقدير لما كان سيحصل في غياب التدخل. المكون الثالث هو التحقيق واستبعاد التفسيرات البديلة. في بعض الحالات، سيكون من الممكن إدراج كل المكونات الثلاث في تقييم الأثر. أما في الحالات المعقدة، فقد لا يكون بالإمكان تقدير الواقع المضاد، وسيحتاج التحليل السببي إلى الاعتماد على المكونات الأخرى.

طرق ممكنة لفحص الواقع (المدى الذي تطابق النتائج الحقيقية فيه ما كان متوقعاً):

دراسات الحالة المقارنة - هل قام التدخل بإحداث النتائج فقط في حالات حينما كانت العناصر الضرورية الأخرى في محلها؟

الاستجابة للجرعة - هل كانت هناك نتائج أفضل للمشاركين الذي تلقوا قدرًا أكبر من التدخل (على سبيل المثال، حضروا المزيد من ورش العمل أو تلقوا المزيد من الدعم)؟

عزو المستفيد/الخبير - هل يعتقد المشاركون/المبلغون أن التدخل قد أحدث فرقاً، وهل يمكنهم تقديم شرح معقول لسبب هذه الحالة؟

التنبؤات - هل تنبأ أولئك المشاركون أو المواقع بتحقيق أفضل الآثار (بسبب جودة التنفيذ و/أو سياق مؤاتٍ)؟ كيف يمكن شرح الحالات الشاذة؟

الزمنية - هل حدثت الآثار في وقت متناغم مع نظرية التغيير - ليس قبل تنفيذ التدخل؟

تشمل الطرق الممكنة لفحص الواقع المضاد (تقدير لما كان سيحصل في غياب التدخل) ما يلي:

الاختلاف في الاختلاف - يُقارن الاختلاف القبلي والبعدي للمجموعة التي تلقت التدخل (حيث لم

مؤشرات التنمية العالمية - يمتلك البنك الدولي بيانات لأكثر من 200 دولة فيما يتعلق بأكثر من 1000 مؤشر¹⁹.

تقدم الملاحظة الإرشادية 3 المزيد من التفاصيل حول الطرق المحددة لقياس أو وصف الآثار، واستخدام الطرق المختلطة (البيانات الكمية والكيفية المستخدمة بطرق تكملية)، وطرق التعامل مع التحديات أثناء قياس أو وصف الآثار.

9. شرح مدى إنتاج التدخل للنتائج الملاحظة

إحدى أهم مميزات تقييم الأثر هي أنه لا يقوم فقط بجمع الأدلة بأن الآثار قد حدثت، بل يحاول أيضاً أن يفهم دور التدخل في إحداثها. ومن النادر أن يكون التدخل هو السبب الوحيد للتغيرات. فالتدخل يعمل عادةً مقترناً مع برامج أخرى أو سياق مؤاتٍ أو عوامل أخرى. وغالباً ما تتعاون مجموعة لإحداث أثر مشترك مثل حينما تشترك المنظمات الدولية غير الحكومية مع الحكومات والمجتمعات المحلية. لذلك، لا يشير "العزو السببي" عادةً إلى العزو الكلي (أي أن التدخل هو السبب الوحيد)، ولكن يشير إلى العزو الجزئي أو إلى تحليل مساهمة التدخل. ويشار إلى ذلك أحياناً بـ "المساهمات المعقولة".

على سبيل المثال، في البحث الزراعي، يمكن أن تكون الآثار المتعلقة بالإنتاجية المحسنة بسبب سلسلة طويلة من البحث وتطوير المنتج والتواصل الأساسي والمطبق. فالاستثمار في أي منها قد يدعي بشكل معقول أنه كان أمراً أساسياً لإحداث الآثار، ولكنه لن يكون قادراً على فعل ذلك بدون التدخلات الأخرى. بعبارة أخرى، يمكن أن يكون تدخلاً ضرورياً ولكن ليس كافياً لإحداث الأثر بنفسه.

يمكن أن يُفيد التحقيق في العزو السببي أو العزو المعقول فيما يتعلق بهذه المكونات. النقطة الأولى هي الواقع - أي مقارنة النتائج الحقيقية بتلك المتوقعة إذا كانت نظرية التغيير

http://www.un.org/esa/sustdev/natlinfo/indicators/methodology_18_sheets.pdf

<http://data.worldbank.org/data-catalog/world-development-indicators> 19

يتم تعيينهم عشوائياً بالاختلاف القبلي والبعدى لأولئك الذين لم يتلقوه.²⁰

الواقع المضاد المبني منطقياً - في بعض الحالات من الموثوق استخدام الخط القاعدي كتقدير للواقع المضاد. على سبيل المثال، حيث تم تركيب مضخة مياه، قد يكون من المعقول قياس الأثر عن طريق مقارنة الوقت المبذول في جلب الماء من مضخة بعيدة قبل وبعد التدخل، بما أنه لا يوجد سبب معقول لأن ينخفض الوقت المستغرق بدون التدخل (وايت، 2007). يمكن أن يدعم تتبع العملية هذا التحليل في كل خطوة من نظرية التغيير.²¹

المقارنات المطابقة - يتطابق كل من المشاركين (الأفراد أو المنظمات أو المجتمعات) مع غير المشاركين على متغيرات يُعتقد أن تكون ذات صلة. ويمكن أن يكون من الصعب مطابقة كل المعايير ذات الصلة بشكل كافٍ.²²

الخطوط القاعدية المتعددة أو الخطوط القاعدية الدوارة - ينظم تنفيذ التدخل عبر الوقت وعدد السكان الذين يؤثر فيهم التدخل. يقوم التحليل بالبحث عن نمط مكرر للتغيير في النتائج المقاسة في كل مجتمع بعد تنفيذ التدخل، بالإضافة إلى غياب التذبذبات الهامة في البيانات في مراحل زمنية أخرى. يُستخدم ذلك بشكل متزايد في التدخلات المتعلقة بالصحة على مستوى السكان.²³

علامات الميول - تخلق هذا التقنية إحصائياً مجموعات مقارنة على أساس تحليل للعوامل التي أثرت في ميول الناس للمشاركة في البرنامج - وهي مفيدة بالأخص حينما تكون المشاركة تطوعية (على سبيل المثال، مشاهدة برنامج تلفزيوني يحتوي على رسائل معززة للصحة).²⁴

التجارب العشوائية المراقبة - يتم تعيين المشاركين المحتملين (أو الجماعات أو الأسر) عشوائياً ليتلقوا التدخل أو ليكونوا في مجموعة مراقبة (إما دون تدخل أو بالتدخل المعتاد) وتقارن النتائج المتوسطة للمجموعات المختلفة.²⁵

انقطاع التراجع - حيثما يتوفر التدخل فقط للمشاركين فوق أو تحت نقطة قطع معينة (مثلاً، الدخل)، يقارن هذا الأسلوب نتائج الأفراد مباشرة تحت نقطة القطع مع أولئك الأفراد فوق نقطة القطع.²⁶

الواقع المضاد المبني إحصائياً - يُستخدم النموذج الإحصائي، مثل تحليل التراجع، لتطوير تقدير لما كان سيحصل في غياب التدخل. ويمكن استخدام ذلك عندما يكون التدخل متناسباً مع واقع الحال - على سبيل المثال، تقييم الأثر لخصخصة خدمات توفير المياه الوطنية.

يمكن أن يكون من الصعب عملياً تطوير واقع مضاد موثوق به. وكثيراً ما يكون من الصعب مطابقة الأفراد أو الجماعات على المتغيرات التي تصنع فرقا حقا. يمكن أن تقوم التجارب العشوائية المراقبة بإنشاء مجموعات غير متكافئة بشكل عشوائي. تعتمد الطرق الأخرى على افتراضات متنوعة قد لا يتم الوفاء بها. وفي حالات التغيير

20 للمزيد من الإطلاع:

Difference-in-differences, http://ec.europa.eu/regional_policy/sources/docgener/evaluation/evalsed/sourcebooks/method_techniques/counterfactual_impact_evaluation/difference-in-differences/difference-in-differences_en.htm.

21 للمزيد من الإطلاع:

Process Tracing in Case Study Research, <http://users.polisci.wisc.edu/kritzer/teaching/ps816/ProcessTracing.htm>.

22 للمزيد من الإطلاع:

Techniques for Improving Constructed Matched Comparison Group Impact/Outcome Evaluation Designs, <http://knol.google.com/k/techniques-for-improving-constructed-matched-comparison-group-impact-outcome>.

23 للمزيد من الإطلاع:

The Multiple Baseline Design for Evaluating Population-Based Research, <http://rds.epi-ucsf.org/ticr/syllabus/courses/23/2011/06/02/Lecture/readings/Hawkins.pdf>.

24 للمزيد من الإطلاع:

Propensity Scores: What Do They Do, How Should I Use Them, and Why Should I Care?, <http://www.chrp.org/love/ASACleveland2003Propensity.pdf>; A Practical Guide to Propensity Score Models, <http://www.academyhealth.org/files/2009/monday/HebertP.pdf>.

25 للمزيد من الإطلاع:

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

26 للمزيد من الإطلاع:

Impact Evaluation: Regression Discontinuity, <http://www.urban.org/toolkit/data-methods/regression.cfm>.

وجد تقييم الأثر لتشريع ارتداء خوذة الدراجة الإجباري أنه كان هناك تراجع كبير في عدد إصابات الرأس بين راكبي الدراجات. وبينما يتناغم ذلك مع نظرية التغيير، فهناك تفسير بديل يقضي بأن مستوى الإصابات قد تراجع بسبب زيادة بناء ممرات الدراجات خلال نفس الفترة. أظهر فحص الإصابات الخطيرة أنه بينما تراجع مستوى إصابات الرأس في تلك الفترة، فإن عدد أنواع الإصابات الأخرى بقي ثابتاً مما يدعم نظرية أن الخوذات هي التي أحدثت التغيير. (والتر وآخرون، 2011)

تشمل بعض الأساليب التي تجمع هذه العناصر المختلفة للتفسير ما يلي:

خطوط ومستويات الأدلة المتعددة - تتم مراجعة مجموعة عريضة من الأدلة من مصادر مختلفة بواسطة هيئة من الخبراء الموثوق بهم الذين يغطون مجموعة من التخصصات ذات الصلة. تحدد الهيئة الاتساق مع نظرية التغيير بينما تحدد وتشرح أيضاً الاستثناءات. يقوم أسلوب خطوط ومستويات الأدلة المتعددة بمراجعة الأدلة من أجل إيجاد علاقة سببية بين التدخل والآثار الملاحظة من حيث قوتها واتساقها وخصوصيتها وزمنيتها وارتباطها مع الأدلة المقبولة الأخرى ومعقوليتها وتشابها الجزئي مع التدخلات المشابهة.²⁹

تحليل المساهمة - أسلوب منهجي يتضمن تطوير نظرية التغيير، ورسم الخرائط للبيانات الموجودة، وتحديد التحديات للنظرية - بما في ذلك الفجوات في الأدلة والروابط السببية المفندة - وجمع تكراري للأدلة الإضافية لتناول ذلك. تقدم الملاحظة الإرشادية بعض المعلومات الإضافية حول تحليل المساهمة.³⁰

رفع التقارير المشتركة بالنتائج - يجمع هذا الأسلوب

السريع والذي لا يمكن التنبؤ به، قد لا يكون من الممكن إنشاء واقع مضاد موثوق به. قد يكون من الممكن بناء حالة تجريبية قوية حيث أحدث التدخل آثاراً معينة، ولكن ليس للتأكد من ما كان سيحصل لو لم يتم تنفيذ التدخل. على سبيل المثال، قد يكون من الممكن إظهار أن تطوير بنية تحتية للمجتمع لتربية الأسماك للاستهلاك أو البيع كان بسبب مشروع محلي بشكل مباشر، بدون القدرة على الإعلان بثقة أن هذا لم يكن ليحصل في غياب المشروع (ربما من خلال مشروع بديل تنفذه منظمة أخرى). ما يستطع تقييم الأثر أن يركز عليه هو العنصرين الآخرين للتحليل السببي - وهما البديلين الواقع والاستبعاد.

المكون الثالث لفهم الارتباطات السببية هو التحقيق واستبعاد التفسيرات البديلة. قد تعكس الآثار الواضحة (أو نقصها) مشاكل منهجية مثل الانحياز في الاختيار (حيث يكون المشاركون مختلفين بشكل منهجي عن غير المشاركين)، وتأثيرات التلوث (حيث ينتفع غير المشاركين من التدخلات أيضاً، مما يقلل الفرق بينهم وبين المشاركين من ناحية الآثار). وقد تعكس الآثار تأثير العوامل الأخرى، بما في ذلك التدخلات الأخرى أو تحركات السكان بين المناطق التي تم تعيينها لتتلقى التدخل وتلك التي لم يتم تعيينها.

تشمل الطرق الممكنة لتحديد واستبعاد التفسيرات الممكنة البديلة ما يلي:

منهج الإقصاء العام - يتم تحديد التفسيرات البديلة الممكنة ومن ثم التحقيق فيها لترى ما إذا كان بالإمكان استبعادها.²⁷

البحث عن الأدلة الداخلة/متابعة الاستثناءات²⁸

²⁹ للمزيد من الإطلاع:

Patricia Rogers, "Matching Impact Evaluation Design to the Nature of the Intervention and the Purpose of the Evaluation", *Journal of Development Effectiveness*, 1 (2009): 217-226. Working paper version available at: http://www.3ieimpact.org/admin/pdfs_papers/50.pdf.

³⁰ للمزيد من الإطلاع:

Contribution Analysis: ILAC Guide Brief 16, http://www.cgia-iac.org/files/publications/briefs/ILAC_Brief16_Contribution_Analysis.pdf; Contribution Analysis, <http://www.scotland.gov.uk/Resource/Doc/175356/0116687.pdf>.

²⁷ للمزيد من الإطلاع:

Can We Infer Causation from Cross-Sectional Data?, http://www7.national-academies.org/BOTA/School-Level%20Data_Michael%20Scriven-Paper.pdf

²⁸ للمزيد من الإطلاع:

Matthew B. Miles and A. Michael Huberman, *Qualitative Data Analysis: An Expanded Sourcebook*. 2nd Edition (Thousand Oaks, California: Sage Publications, 1994).

متفق عليه مع مستخدمين مقصودين يمكن عندها أن يتم استخدامهم في صنع الأدلة بشفافية. ويتضمن المقياس عنواناً لكل نقطة (على سبيل المثال، "غير ناجح"، "ناجح إلى حد ما"، "ناجح جداً") ووصف لما تبدو عليه كل منها.³³

11. رفع التقارير بالنتائج ودعم استخدامها

ينبغي الاتفاق على شكل تقرير التقييم عند تخطيط تقييم الأثر. بعض المنظمات لديها أشكال تقرير قياسية، وتشمل متطلبات الطول، التي يجب إتباعها. وفي الحالات الأخرى، من المهم الاتفاق على "تقرير هيكل" للعناوين الرئيسية والفرعية قبل كتابة التقرير.

يكون الوصول إلى تقارير التقييم أكثر سهولة حينما تكون منظمة حول أسئلة التقييم الرئيسية، بدلاً من رفع التقارير بشكل منفصل حول البيانات من المكونات المختلفة لجمع البيانات.³⁴

يمكن تعزيز جودة تقارير التقييم عن طريق إشراك الملائم لأصحاب المصالح. وحتى حينما يتولى التقييم مقيّم خارجي مستقل، يمكن إشراك أصحاب المصالح عن طريق تقديم استجابات رسمية للنتائج والتعليق على البيانات وكيف تم تفسيرها.

وحيثما يتم تضمين التوصيات في تقارير التقييم، يجب أن تُدعم بالأدلة من نتائج التقييم وتُطور حول ملامحة وعملية التوصيات. لا يمكن أن يحسّن إشراك أصحاب المصالح المناسبين في تطوير التوصيات الجانب العملي للتوصيات فحسب، بل يمكن أيضاً أن يؤدي إلى امتلاك أصحاب المصالح لها والالتزام بتطبيقها.

يمكن أن يُعزّز استخدام تقارير تقييم الأثر عن طريق أشكال رفع التقارير المبتكرة، والعروض التقديمية الشفهية،

الجديد بين تحليل المساهمة وخطوط ومستويات الأدلة المتعددة. فهو يقوم على رسم الخرائط للبيانات الموجودة مقابل نظرية التغيير وسد الفجوات المهمة في الأدلة من خلال جمع البيانات الإضافية المستهدفة. ثم يتم استخدام مزيج من مراجعات الخبراء والاستشارات المجتمعية للتحقق من مصداقية الأدلة فيما يتعلق بالآثار التي حدثت والمدى الذي يمكن أن نعزو هذه الآثار بشكل معقول إلى التدخل.³¹

لم يستطع تقييم برنامج تنمية تنفيذي عبر الحكومة أن يستخدم مجموعة مراقبة عشوائية، لأن تعيين الأشخاص عشوائياً ليكونوا في مجموعة مراقبة - أو حتى المشاركة في البرنامج - كان أمراً مستحيلاً. ولم يستطع التقييم كذلك أن يستخدم مجموعة المراقبة، بسبب أن طبيعة البرنامج كانت تقتضي أن الأشخاص المقبولين فيه كانوا مختلفين بشكل منهجي عن الأشخاص الذين لم يُقبلوا. بدلاً من ذلك، استخدم التقييم استراتيجيات التفسير السببي، بما في ذلك العزو من قبل المنتفعين والزمنية والخصوصية (كانت التغييرات في المناطق المحددة التي تناولها البرنامج). (دايفدسون، 2006)

10. صنع الأدلة

من النادر أن نبنى الحكم الكلي للتقييم على أساس قياس أداء واحد. فعادةً يتطلب الأمر صنع الأدلة بشأن الأداء عبر الأبعاد المختلفة.

والطريقة المعتادة للقيام بذلك هي تطوير مقياس ذو وزن حيث تُبنى نسبة تقدير الأداء الكلي فيه على أساس كل بعد من الأبعاد. ولكن كثيراً ما يكون للمقياس العددي ذو الوزن مشاكله، بما في ذلك الأوزان الاعتباطية ونقص الانتباه للعناصر الأساسية.³²

طريقة بديلة هي تطوير مقياس (أو قاعدة) تقدير عالمي

³³ للمزيد من الإطلاع:

The Rubric Revolution, <http://kinnect.co.nz/wp-content/uploads/2011/09/AES-2011-Rubric-Revolution-Davidson-Wehipeihana-McKegg-xx.pdf>.

³⁴

E. Jane Davidson, "Improving Evaluation Questions and Answers: Getting Actionable Answers for Real-World Decision Makers" (demonstration session at the 2009 Conference of the American Evaluation Association, Orlando, FL, November 18, 2009), <<http://comm.eval.org/resources/view/document/?DocumentKey=e5bac388-f1e6-45ab-9e78-10e60ceao666>>.

³¹ للمزيد من الإطلاع:

Collaborative Outcomes Reporting Technique, <http://www.aes.asn.au/regions/vic/2010/COR%20AES%202010.pdf>.

³² للمزيد من الإطلاع:

The Synthesis Problem: Issues and Methods in the Combination of Evaluation Results Into Overall Evaluative Conclusions, <http://realevaluation.com/pres/SynthHNL.pdf>.

وفرص المشاركة مع الآخرين ومناقشة المعاني الضمنية لتقييم الأثر، وعن طريق ضمان استمرار قابلية الوصول إلى التقارير بالنسبة للمستخدمين المحتملين.

تتناول الملاحظة الإرشادية 4 الحاجة إلى توصيل النتائج بشكل جيد للجمهور المقصود.

طوال فترة التدخل، ويشمل ذلك بيانات الخط القاعدي، ويسمح بخيار استخدام طرق مثل التجارب العشوائية المراقبة، والتي تتطلب إنشاء مجموعة مراقبة معينة عشوائياً من بداية التنفيذ.

13. ما هو المطلوب لإجراء تقييم أثر ذي جودة؟

من الممكن أن يفيد التفكير بشأن التقييم ذي الجودة من ناحية الضرورات الأساسية الخمسة المتنافسة: الفائدة والدقة والأخلاق والتطبيق العملي والمسؤولية.³⁵ وكثيراً ما تكون هذه المعايير الخمسة في توتر - على سبيل المثال، تقييم الأثر الأكثر شمولية والذي سيكون أكثر دقة قد لا يكون عملياً من ناحية الموارد المتوفرة، أو يكون متطفلاً في البيانات التي يجمعها، أو قد يستغرق وقتاً طويلاً جداً ليكتمل من أجل أن يشكل القرارات الرئيسية حول مستقبل التدخل.

الفائدة - تقييم الأثر الجيد يكون مفيداً. يمكن تعزيز الفائدة المحتملة من التقييم عن طريق التخطيط لكيفية استخدامه من البداية، ويشمل ذلك ربطه بعمليات وتوقيت اتخاذ القرارات التنظيمية، وأن يكون واضحاً من حيث سبب إجراءاته، ومن الذي سيستخدمه، وإشراك أصحاب المصالح الرئيسيين في العملية، ومن ثم اختيار التصاميم والطرق لتحقيق هذا الغرض.

الدقة - تقييم الأثر الجيد يكون دقيقاً. فهو يعطي انتباهاً إلى جميع الأوجه المهمة، ويلاحظ ما إذا كان أي منها غير مقصود. ويعطي انتباهاً إلى توزيع الآثار، ويلاحظ إذا كان بعض الأشخاص فقط قد استفادوا ومن هم أولئك الأشخاص. تتطلب الدقة استخدام الأدلة الملائمة، ويشمل ذلك البيانات الكمية والكيفية، والتفسيرات الملائمة، والشفافية بشأن مصادر البيانات التي تم استخدامها وقيودها. وتشمل إستراتيجيات تحقيق الدقة أنظمة التحقق من جودة البيانات في مرحلة جمعها وخلال معالجتها، وبأنه قد تم رفع التقارير بشأن النتائج بنزاهة وشمولية ووضوح.

12. متى يجب إجراء تقييم الأثر؟

ينبغي أن يُشرع في عمل تقييم الأثر حينما توجد الحاجة الواضحة والنية في استخدام النتائج. فلو كانت كل التدخلات تتطلب عمل تقييم للأثر، فسيخاطر المقيّمون بإحدى أمرين، إما طلب الكثير من الموارد أو نشر هذه الموارد بشكل كبير بحيث تجعل التقييمات سطحية. تتمثل الإستراتيجية الأكثر فعالية في تركيز موارد تقييم الأثر على التدخلات التي تكون فيها الفائدة أكثر ترجيحاً:

- التدخلات المبتكرة والبرامج التجريبية التي يمكن أن تتم زيادتها أو تكرارها، إذا أثبتت أنها ناجحة.
- التدخلات التي لا تُفهم آثارها بشكل جيد، والحاجة لأدلة أفضل لتشكيل القرارات حول الاستمرار في تمويلها أو إعادة توجيه التمويل إلى التدخلات الأخرى.
- التقييمات الدورية لأثر مجموعة من التدخلات في قطاع أو منطقة لتوجيه السياسة وتصميم التدخل المستقبلي وقرارات التمويل.
- التدخلات ذات الخطورة العالية مثل الاستثمار الكبير (الحالي أو المستقبلي)، أو احتمالية عالية لحدوث آثار سلبية كبيرة أو مشكلات سياسية حساسة.
- التدخلات حيث توجد حاجة لأصحاب المصالح لفهم رؤى ومساهمات كل منهم بشكل أفضل.

توقيت تقييم الأثر مهم. فلو تم إجراء التقييم مبكراً جداً، قد لا يكون هناك أدلة كافية على حدوث الآثار أو استدامتها. ولو تم إجراؤه متأخراً جداً، فقد يكون من الصعب المتابعة مع المشاركين ويكون الوقت قد تأخر على التأثير في القرارات المتعلقة بالمسار المستقبلي للتدخل. وعلى أي حال، من الأفضل تخطيط تقييم الأثر من بداية التدخل حيثما كان ذلك ممكناً. فذلك يسمح بجمع الأدلة

35

Joint Committee Standards for Educational Evaluation <http://www.jcsee.org/program-evaluation-standards/program-evaluation-standards-statements>.

تم تطوير هذه الفاييس في الأصل للتقييم التعليمي ولكنها غالباً ما تستخدم بشكل كبير على نطاق أوسع.

الآداب (الأخلاق) - يجب أن يتم تناول المسائل الأخلاقية بشكل وافي - بما في ذلك الخصوصية والسرية، بالإضافة إلى الآثار الضارة المحتملة من المشاركة في التقييم. بعض المسائل الأخلاقية، كالحاجة إلى حفظ الوعود التي قُطعت حول السرية والخصوصية، تعد شائعة عبر الأنواع المختلفة من التقييمات والبحث. وهناك مسائل أخرى خاصة بتقييم الأثر. يُثار القلق أحياناً حول الأخلاقيات وراء استخدام تصميم التجارب العشوائية المراقبة، حيث أنها تتضمن منع التدخل عن بعض الأشخاص (المجموعة المراقبة). ويقل القلق المتعلق بالأخلاقيات حين يُوزع الوصول إلى التدخل في أية حالة، ويمكن التعامل مع ذلك عن طريق تعيين المجموعة المراقبة في طابور لكي يتلقوا التدخل بعد الانتهاء من تقييم المرحلة الأولى (إذا أظهرت أنها فعالة). ولكن هذه الإستراتيجية ممكنة فقط حينما تكون الآثار (أو تنبؤاتها الموثوق بها) واضحة مبكراً، وحين يبقى التدخل مناسباً للمجموعة المراقبة في وقت انتهاء التقييم.

وتوجد أيضاً مسائل أخلاقية محتملة من ناحية من تُلبي مصالحه عبر التقييم. تناقش مؤسسة التقييم الأمريكية هذا الأمر في المبادئ الإرشادية الخاصة بها من ناحية "مسؤوليات الرافهية الشعبية والعامة":

يبين المقيّمون وبأخذون في الاعتبار تنوع القيم والمصالح الشعبية والعامة، وبالتالي يجب عليهم أن:

1. يشملوا الرؤى والمصالح ذات العلاقة الخاصة بالمجموعة الكاملة من أصحاب المصالح.
2. عدم وضع النتائج والعمليات المباشرة للتقييم في الاعتبار فحسب، بل أيضاً الافتراضات الأوسع والتضمينات والآثار الجانبية المحتملة.
3. السماح لأصحاب المصالح بالوصول إلى المعلومات التقييمية ونشرها بشكل نشط، وعرض نتائج التقييم في أشكال يمكن فهمها تحترم الأشخاص وتحفظ الوعود المتعلقة بالخصوصية.
4. الحفاظ على التوازن بين حاجات ومصالح العميل وصاحب المصلحة الآخر.
5. وضع خير ومصصلحة الجمهور في الاعتبار، والمضي إلى أبعد من تحليل مصالح صاحب

مصلحة معين وذلك من أجل اعتبار رافهية المجتمع بأكمله.

من الضروري عادة الحصول على الموافقة الرسمية من لجنة المراجعة المؤسسية المناسبة للبدء في تقييم الأثر. وتحتاج طلبات الموافقة إلى إتباع النمط الموصوف وتناول مسائل الإحسان والعدل والاحترام.³⁶

العملية - يجب أن يكون تقييم الأثر عملياً. ويجب أن يأخذ في الاعتبار الموارد المتوفرة (الوقت والمال والخبرة والبيانات الموجودة) وتوقيت الحاجة إلى نتائج التقييم لتشكيل القرارات. يمكن أن تقوم الشركة مع واحد أو أكثر من المهنيين أو مؤسسات البحث أو الجامعات أو مؤسسات المجتمع المدني بتوفير الموارد الضرورية.

المساءلة - يجب أن توضح التقييمات الأدلة والمعايير التي قامت عليها الاستنتاجات وتعترف بقيودها. الشفافية في البيانات أمر مهم، ويشمل ذلك إظهار المصادر التي استخدمت لأي سؤال من أسئلة التقييم. العملية الرسمية لما بعد التقييم - وهي أن تجعل تقييمك الخاص يُقيم عن طريق الموافقة على خطة التقييم ومن ثم تقرير التقييم - بواسطة مراجع خبير أو لجنة من الأفراد الذين يتمتعون باستقلالية ونزاهة محترمة يمكنها أن تحسن المساءلة المتعلقة بتقييم الأثر.

14. التحديات الشائعة في تطوير تقييم الأثر

تم وصف عدد من التحديات الشائعة للتقييم التطويري أدناه بالإضافة إلى بعض الاقتراحات حول كيفية التعامل معها.

• **التغير في التنفيذ وفي البيئة عبر المواقع المختلفة**
قد يكون التدخل تم تنفيذه بطرق مختلفة تماماً ليناسب سياقات مختلفة في مكاتب دول مختلفة حول العالم أو في أماكن جغرافية مختلفة داخل الدولة. من الممكن أن تفيد مقارنة نظريات التغير لكل موقع. وبالأخص تحديد ما إذا كانت

36 للمزيد من الاطلاع:

Evaluation Consent and the Institutional Review Board Process, <http://meera.snre.umich.edu/plan-an-evaluation/plonearticlemultitpage.200730.3630902539-10-/evaluation-consent-and-the-institutional-review-board-process>.

المواقع المختلفة تستخدم نفس النظرية حول كيفية حدوث التغيير (مثلاً عن طريق زيادة وعي الأشخاص باستحقاقاتهم للخدمات) ولكن تستخدم نظريات عمل مختلفة (مثلاً المنشورات المطبوعة مقابل مسرح المجتمع)، أو ما إذا كانت تستخدم نظريات تغيير مختلفة تماماً (مثلاً زيادة وعي الأشخاص باستحقاقاتهم للخدمات في موقع واحد مقابل تقليل حواجز الوصول للخدمات - مثل رسوم المستخدم - من خلال المناصرة في موقع آخر).

• الآثار المتغيرة

كثيراً ما تعمل تدخلات التطوير بشكل جيد فقط لبعض الأشخاص، وقد تكون غير فعالة أو حتى مضرّة لبعض الأشخاص الآخرين. بالإضافة إلى ذلك، كثيراً ما يتأثر نجاح التدخل من حيث تحقيق الآثار المرغوبة بجودة التنفيذ. لذلك من المهم ألا يتم حساب التأثير المتوسط ورفع التقرير عنه فحسب، بل أيضاً التحقق من التأثيرات التباينية. وهذا يتطلب جمع الأدلة على جودة التنفيذ والبيانات حول عوامل السياق التي قد تؤثر على الآثار حيثما أمكن ذلك، بما في ذلك شخصيات المشاركين وبيئة التنفيذ.

• المكونات المتنوعة

قد يشمل البرنامج مجموعة متنوعة من المشاريع، وبرغم ذلك من الضروري عمل تقييم كلي لأثر البرنامج بأكمله. من الممكن أن يفيد تطوير نظرية تغيير كلية للبرنامج وجمع المكونات المختلفة. وأحياناً يكون من الممكن القيام بذلك في مرحلة التخطيط، إلا أنه قد يكون من الضروري عمل ذلك بأثر رجعي، خصوصاً حيثما تكون المشاريع والمكونات قد تغيرت مع الوقت.

• المقاييس على المدى الطويل

في أحيان كثيرة، لن تكون الآثار المقصودة واضحة لسنوات عديدة، ولكن الأدلة المطلوبة لتشكيل القرارات قبل ذلك الوقت (مثلاً مسألة ما إذا كان سيتم إطلاق مرحلة تالية أو تكرار النموذج

في مكان آخر). يمكن أن تحدد نظرية التغيير النتائج المتوسطة التي قد تكون واضحة في عمر التقييم. في بعض الحالات، يمكن استخدام أدلة البحث ملء الروابط اللاحقة، ولتقدير الآثار المحتملة على افتراض تحقيق النتائج المتوسطة. وينبغي أيضاً أن يتم وضع المسار المتوقع للتغيير في الاعتبار - حين يُرجح أن تكون الآثار واضحة.³⁷

• تأثير البرامج والعوامل الأخرى

إن آثار تدخلات التطوير تتأثر بشكل كبير بأنشطة البرامج وعوامل السياق الأخرى التي قد تدعم أو تعوق الآثار الجارية تحقيقها. على سبيل المثال، ستؤدي تحويلات النقد المشروطة بحضور المدرسة إلى إنجاز دراسي محسّن للطلاب فقط في الحالات التي تقوم المدارس فيها بتدريس الطلاب بشكل ملائم. من الممكن تحديد هذه البرامج وعوامل السياق الأخرى كجزء من تطوير نظرية التغيير وجمع الأدلة حولها والبحث عن الأمهات في البيانات.

• القيود على الموارد

قد تحتوي الأدلة الموجودة (في شكل وثائق البرنامج وبيانات الخط القاعدي والإحصائيات الرسمية) على فجوات، وقد يكون هناك موارد قليلة (من حيث التمويل أو وقت الطاقم أو الوصول إلى الخبرة الفنية المتخصصة) لجمع أنواع الأدلة المطلوبة لإجراء تقييم أثر ذي جودة. ولإجراء تقييم معين، حينما تكون الأدلة الموجودة قليلة مع قلة الموارد لجمع أدلة إضافية، قد تقدم المقابلات الرئيسية من المبلغين المتنوعين بيانات كافية، بما في ذلك إعادة إنشاء بيانات الخط القاعدي. ويمكن أن يقلل التخطيط المسبق لتقييم الأثر القيود على الموارد عن طريق بناء موارد كافية في مرحلة التصميم ووضع الميزانية و/أو تقسيم موارد التقييم عبر التدخلات لكي تركز أكثر على عدد صغير

37 للمزيد من الإطلاع:

Michael Woolcock on The Importance of Time and Trajectories in Understanding Project Effectiveness, <http://blogs.worldbank.org/impactevaluations/guest-post-michael-woolcock-on-the-importance-of-time-and-trajectories-in-understanding-project-effe>.

Association, Portland, Ore., Nov. 2006 <http://realevaluation.com/pres/causation-anzea09.pdf>.

- Davidson, E. J. (2009) "Improving Evaluation Questions and Answers: Getting Actionable Answers for Real-World Decision Makers." Demonstration session at the 2009 Conference of the American Evaluation Association, Orlando, Fla., Nov. 2009. <http://comm.eval.org/resources/vi/ewdocument/?DocumentKey=e5bac388-f1e6-45ab-9e78-10e60cea0666>.
- Funnell S. and Rogers, P. (2011) *Purposeful Program Theory: Effective Use of Theories of Change and Logic Models*. San Francisco: Jossey-Bass/Wiley.
- Guijt, I. (1999) Participatory Monitoring and Evaluation for Natural Resource Management and Research. Socio-economic Methodologies for Natural Resources Research. Chatham, UK: Natural Resources Institute. <http://www.nri.org/publications/bpg/bpgo4.pdf>
- Miles, M. and Huberman, M. (1994) *Qualitative Data Analysis: An Expanded Sourcebook* (2nd ed.) Thousand Oaks California; Sage Publications. *Outlines strategies for checking causal explanations, including searching for disconfirming evidence, following up exceptions, making and testing predictions.*
- Patton, MQ (2008) "State of the Art in Measuring Development Assistance." *Discusses the importance of interpretation and managing uncertainty in effective management.*
- Paz R., Dorward A., Douthwaite B. (2006). "Methodological Guide for Evaluation of

من التقييمات الأكثر شمولاً للتدخلات المهمة استراتيجياً.

الخلاصة

يجب أن يبدأ تقييم الأثر بخطة توضح الأغراض المقصودة، وتشمل تحديد المستخدمين المقصودين، وأسئلة التقييم الرئيسية التي هي معدة للإجابة عليها، وكيف ستتناول المكونات الستة لتقييم الأثر - وهي توضيح القيم، وتطوير نظرية التغيير، وقياس أو وصف المتغيرات المهمة، وشرح الأمور التي أنتجت الآثار، وصنع الأدلة، ورفع التقارير ودعم الاستخدام. والقيام بمراجعة هذه الخطة (بواسطة أشخاص بمن فيهم المستخدمين المقصودين) سيزيد من احتمالية إنتاج تقييم أثر عالي الجودة يُستخدم بالفعل.

المراجع والموارد المفيدة الأخرى

- Alton-Lee, A. (2003) "Quality Teaching for Diverse Students in Schooling: Best Evidence Synthesis." Wellington, New Zealand: Ministry of Education. <http://www.educationcounts.govt.nz/publications/series/2515/5959> *An example of synthesizing evidence from diverse sources to understand what works for whom.*
- Catley A., Burns, J., Abebe, D., Sufi, O. Participatory Impact Assessment: A Guide for Practitioners. Boston: Tufts University. <http://www.preventionweb.net/english/professional/publications/v.php?id=9679>
- Chambers R. (2007) "Who Counts? The Quiet Revolution of Participation and Numbers" Working Paper No. 296, Brighton: Institute of Development Studies. <http://www.ids.ac.uk/files/Wp296.pdf>
- Davidson, E. J. (2006) "Causal Inference Nuts and Bolts." Demonstration session at the 2006 Conference of the American Evaluation

Pro-Poor Impact of Small-Scale Agricultural Projects.” Centre for Development and Poverty Reduction. Imperial College, London. <http://boru.pbworks.com/f/modulosjano7.pdf>
Describes 22 methods and tools that can be used to evaluate the direct and indirect impacts of innovation adoption.

Roche, C. (1999) Impact Assessment for Development Agencies: Learning to Value Change Oxford: OXFAM, Novib

Rogers, Patricia J. (2009) “Matching Impact Evaluation Design to the Nature of the Intervention and the Purpose of the Evaluation”, *Journal of Development Effectiveness*, 1(3): 217-226. Working paper version available at: http://www.3ieimpact.org/admin/pdfs_papers/50.pdf.

Walter, S., Olivier, J., Churches, T. and Grzebieta, R. (2011). “The impact of compulsory cycle helmet legislation on cyclist helmet head injuries in New South Wales, Australia”, *Accident Analysis and Prevention*, 43 : 2064-2071.

White, S. and J. Petit (2004) Participatory Methods and the Measurement of Wellbeing Participatory Learning and Action 50, London: IIED

www.betterevaluation.org – information on evaluation methods for development, including user-contributed examples and comments

www.mymande.org – information, videos and links to information about evaluation methods