

Document:	EB 2022/135/R.29
Agenda:	16(a)
Date:	15 March 2022
Distribution:	Public
Original:	English

A



دليل التقييم المنقح في الصندوق - الجزء الأول

مذكرة إلى ممثلي الدول الأعضاء في المجلس التنفيذي

الأشخاص المرجعيون:

نشر الوثائق:

Deirdre Mc Grenra

مديرة مكتب الحوكمة المؤسسية

والعلاقات مع الدول الأعضاء

رقم الهاتف: +39 06 5459 2374

البريد الإلكتروني: gb@ifad.org

الأسئلة التقنية:

Indran A. Naidoo

مدير مكتب التقييم المستقل في الصندوق

رقم الهاتف: +39 06 5459 2274

البريد الإلكتروني: i.naidoo@ifad.org

Fabrizio Felloni

نائب مدير

مكتب التقييم المستقل في الصندوق

رقم الهاتف: +39 06 5459 2361

البريد الإلكتروني: f.felloni@ifad.org

Donal Brown

نائب الرئيس المساعد

دائرة إدارة البرامج

رقم الهاتف: +39 06 5459 2448

البريد الإلكتروني: d.brown@ifad.org

Nigel Brett

مدير

شعبة سياسات العمليات والنتائج

رقم الهاتف: +39 06 5459 2516

البريد الإلكتروني: n.brett@ifad.org

المجلس التنفيذي – الدورة الخامسة والثلاثون بعد المائة

روما، 25-27 أبريل/نيسان 2022

للعلم

جدول المحتويات

ii

شكر وتقدير

iii

موجز تنفيذي

شكر وتقدير

أعدت هذا الدليل فرقة عمل مؤلفة من موظفي مكتب التقييم المستقل في الصندوق، وشعبة سياسات العمليات والنتائج، وشعبة البحوث وتقدير الأثر، بتوجيه عام من: Indran Naidoo، مدير مكتب التقييم المستقل في الصندوق؛ وDonal Brown، نائب الرئيس المساعد، دائرة إدارة البرامج؛ وMeike van Ginnerken، نائبة الرئيس المساعدة السابقة، دائرة الاستراتيجية وإدارة المعرفة؛ ثم مؤخرًا Jyotsna Puri، نائبة الرئيس المساعدة، دائرة الاستراتيجية وإدارة المعرفة. وتولى رئاسة فرقة العمل Fabrizio Felloni، نائب مدير مكتب التقييم المستقل في الصندوق؛ وNigel Brett، مدير شعبة سياسات العمليات والنتائج؛ وSara Savastano، مديرة شعبة البحوث وتقدير الأثر. وشارك في رئاسة فرقة العمل في مراحلها الأولى Thomas Eriksson، المدير السابق لشعبة سياسات العمليات والنتائج. وشمل الأعضاء الآخرون في فرقة العمل: Simona Somma، موظفة التقييم في مكتب التقييم المستقل في الصندوق، التي أشرفت على إدارة العملية؛ وLauren Phillips، مستشارة رئيسية سابقة في شعبة سياسات العمليات والنتائج؛ وAslihan Arslan، كبيرة الخبراء الاقتصاديين في شعبة البحوث وتقدير الأثر؛ وRomina Cavatassi، خبيرة اقتصادية رئيسية في شعبة البحوث وتقدير الأثر؛ وDimitra Stamatopoulos، أخصائية السياسات والنتائج في شعبة سياسات العمليات والنتائج؛ وXiaozhe Zhang، أخصائية السياسات والنتائج في شعبة سياسات العمليات والنتائج؛ وRodney Cooke، استشاري في شعبة سياسات العمليات والنتائج؛ وEmanuele Zucchini، استشاري في شعبة البحوث وتقدير الأثر.

وقدمت Paula Silva Villanueva، كبيرة الاستشاريين، مشورة تقنية موسّعة في صياغة الجزء الأول من الدليل. وقدمت Dorothy Lucks، كبيرة الاستشاريين، تعليقات على الجزء الأول وأسدت مشورتها إلى مكتب التقييم المستقل في إعداد فصول الجزء الثاني ذات الصلة.

واستفاد التقرير من الاستعراض الموضوعي الذي أجراه الفريق الاستشاري للتقييم التابع لمكتب التقييم المستقل في الصندوق والمؤلف من: Bagele Chilisa، جامعة بوتسوانا؛ وGonzalo Hernández Licon، جامعة أكسفورد؛ وHans Lundgren، الرئيس السابق لأمانة التقييم، شبكة التقييم في لجنة المساعدة الإنمائية التابعة لمنظمة التعاون والتنمية في الميدان الاقتصادي؛ وDonna Mertens، أستاذة فخريّة في جامعة Gallaudet؛ وRob van den Berg، أستاذ زائر، جامعة كينغز كوليدج في لندن والمدير السابق لمكتب التقييم المستقل في مرفق البيئة العالمية.

ونوقشت في يونيو/حزيران 2021 مسودة نسخة من الدليل مع فريق خبراء دولي بترشيح مشترك من مكتب التقييم المستقل وإدارة الصندوق. وتألّف فريق الخبراء الدولي من: Marie Gaarder، المديرة التنفيذية للمبادرة الدولية لتقييم الأثر؛ وStefano Gagliarducci، أستاذ في قسم الاقتصاد والمالية في جامعة روما تور فيرغاتا؛ وMegan Kennedy-Chouane، رئيسة وحدة التقييم في إدارة التعاون الإنمائي في لجنة المساعدة الإنمائية التابعة لمنظمة التعاون والتنمية في الميدان الاقتصادي؛ وSebastian Stolorz، كبير موظفي العمليات، إدارة الاستراتيجية والمخاطر والتعلم في البنك الدولي؛ وMaya Vijayaraghavan، مستشارة أساليب منهجية، إدارة التقييم المستقل في مصرف التنمية الآسيوي. وبالإضافة إلى هؤلاء، قدمت Andrea Cook، مديرة التقييم في برنامج الأغذية العالمي تعليقات مكتوبة.

والشكر موصول أيضا لكثير من الموظفين في مكتب التقييم المستقل، ودائرة إدارة البرامج، ودائرة الاستراتيجية وإدارة المعرفة على ما قدموه من إسهامات.

موجز تنفيذي

أولا- الخلفية

- 1- أوصى استعراض الأقران الخارجيين لوظيفة التقييم في الصندوق لعام 2019 بأن يُعد مكتب التقييم المستقل في الصندوق وإدارة الصندوق سياسة تقييم معدلة، وبأن يضع مكتب التقييم المستقل استراتيجية متعددة السنوات للتقييم المستقل، وبأن يضع المكتب دليلاً منقحاً للتقييم بالتعاون مع إدارة الصندوق.
- 2- وفي أبريل/نيسان 2021، وافق المجلس التنفيذي على سياسة التقييم المعدلة. وفي ديسمبر/كانون الأول 2021، وافق المجلس أيضاً على استراتيجية التقييم المستقل المتعددة السنوات لمكتب التقييم المستقل، وإطار الفعالية الإنمائية المحدّث للإدارة الذي شمل استراتيجية الإدارة للتقييم الذاتي.
- 3- وبموازاة هذه الجهود، أعدت فرقة عمل مؤلفة من موظفين من مكتب التقييم المستقل وشعبة سياسات العمليات والنتائج (دائرة إدارة البرامج) وشعبة البحوث وتقدير الأثر (دائرة الاستراتيجية وإدارة المعرفة) دليلاً جديداً للتقييم. وتتوافق الوثيقة (انظر الذيل) مع الجزء الأول من الدليل. ولا يزال العمل جارياً في إعداد الجزء الثاني كما هو موضح أدناه.

ثانياً- السمات الرئيسية لإعداد الدليل

- 4- بينما لم يتناول الإصداران السابقان من دليل التقييم (2009 و2015) سوى التقييم المستقل، ينطبق دليل التقييم الجديد لأول مرة على التقييم الذاتي والتقييم المستقل ويُقدّم إرشادات منهجية ومعايير للتقييم على نطاق المنظمة. وستُساعد المعايير المشتركة على تعزيز جودة التقييمين الذاتي والمستقل والحفاظ على الاتساق بينهما في نفس الوقت الذي ستُعزز فيه أيضاً معايير الرقابة في الصندوق. وستعزز في نهاية المطاف ثقافة أقوى للنتائج والتقييم في الصندوق.
- 5- ويتبع دليل التقييم الجديد الممارسات والمعايير الدولية، ولا سيما تلك التي وضعها فريق الأمم المتحدة المعني بالتقييم، وفريق التعاون في مجال التقييم التابع للمصارف الإنمائية المتعددة الأطراف، وشبكة التقييم في لجنة المساعدة الإنمائية التابعة لمنظمة التعاون والتنمية في الميدان الاقتصادي، بما في ذلك نسختها الأخيرة من معايير التقييم.
- 6- وسيُتاح الدليل في منصة إلكترونية، مما سيُسهل إجراء أي تحديثات وتنقيحات. وسيُألف الدليل من جزأين. ويحتوي الجزء الأول على المراجع والمعايير المشتركة بين مكتب التقييم المستقل وإدارة الصندوق، بينما يشمل الجزء الثاني إرشادات بشأن المنهجية والشكل والعملية المتبعة في منتجات التقييم المحددة الصادرة عن مكتب التقييم المستقل وإدارة الصندوق. وسيجري الربط بين الجزأين الأول والثاني لضمان اتساق محتوياتهما وتعزيز كل منهما الآخر.
- 7- وسيكون المستخدمون الرئيسيون لهذا الدليل هم موظفو الصندوق والاستشاريون الذين يديرون عمليات التقييم أو يشاركون في عمليات التقييم المستقل والذاتي. ويشمل الجمهور الثانوي أصحاب المصلحة والشركاء في الصندوق، مثل الحكومات والقطاع الخاص والمجتمع المدني والشركاء الإنمائيين المشاركين في عمليات التخطيط والرصد والتقييم في الصندوق. ويمكن أن يكون الدليل أيضاً موضع اهتمام لأشخاص خارج الصندوق، مثل الممارسين في مجال التنمية الريفية.
- 8- ويستفيد الدليل من استعراض وتعليقات الخبراء الدوليين، ولا سيما الفريق الاستشاري للتقييم التابع لمكتب التقييم المستقل والمؤلف من: Bagele Chilisa، جامعة بوتسوانا؛ وGonzalo Hernández Licona، جامعة أكسفورد؛ وHans Lundgren، الرئيس السابق لأمانة التقييم، شبكة التقييم في لجنة المساعدة الإنمائية

التابعة لمنظمة التعاون والتنمية في الميدان الاقتصادي؛ وDonna Mertens، أستاذة فخريّة، جامعة Gallaudet؛ وRob Van den Berg، أستاذ زائر، جامعة كينغز كوليدج لندن، والمدير السابق لمكتب التقييم المستقل في مرفق البيئة العالمية.

9- وبالإضافة إلى ذلك، عيّن مكتب التقييم المستقل في الصندوق وإدارة الصندوق في يونيو/حزيران 2021 فريقاً دولياً من الخبراء للتعبير على النسخة الأولى من الدليل. وتألّف الفريق من: Marie Gaarder، المديرّة التنفيذيّة للمبادرة الدوليّة لتقييم الأثر؛ وStefano Gagliarducci، أستاذ في قسم الاقتصاد والمالية، جامعة روما تور فيرغاتا؛ وMegan Kennedy-Chouane، رئيسة وحدة التقييم في إدارة التعاون الإنمائي، لجنة المساعدة الإنمائيّة التابعة لمنظمة التعاون والتنمية في الميدان الاقتصادي؛ وSebastian Stolorz، كبير موظفي العمليات في إدارة الاستراتيجية والمخاطر والتعلم في البنك الدولي؛ وMaya Vijayaraghavan، مستشارة أساليب منهجية، إدارة التقييم المستقل في مصرف التنمية الآسيوي. وبالإضافة إلى ذلك، قدمت Andrea Cook، مديرة التقييم في برنامج الأغذية العالمي، تعليقات مكتوبة.

ثالثاً- المحتويات الرئيسيّة في الجزء الأول من الدليل وما استجد فيه بالمقارنة مع النسخة السابقة

10- يتضمن الجزء الأول الذي يتناول أسس التقييم في الصندوق، ثلاثة فصول. ويُحدد الفصل الأول، وهو عبارة عن مقدّمة، جهود الصندوق في مجال التقييم في سياق عمله من أجل تحقيق تحول ريفي شامل ومستدام، وخطة التنمية المستدامة لعام 2030. ويعرض الفصل الأول أيضاً آخر المعايير الأخلاقية للتقييم التي وضعها فريق الأمم المتحدة المعني بالتقييم. ويعرض الفصل الثاني نُظم التقييم والتعلم على نطاق الصندوق ومسارات التعاون بين التقييميين الذاتي والمستقل وكذلك سُبل ترجمة نتائج التقييم إلى معرفة ودروس مستفادة. ويوضح الفصل الثالث الأسس المنهجية المنطبقة على جميع التقييمات، بدءاً من نُهج تحديد النطاق والتصميم وصولاً إلى معايير التقييم وتصنيفات مختلف أنواع الأداء، وكذلك أساليب جمع البيانات وتحليلها.

11- ويستند الدليل إلى المؤلفات المعاصرة عن التقييم وإلى خبرة مكتب التقييم المستقل وإدارة الصندوق، فضلاً عن التطورات في ممارسات التقييم منذ إطلاق خطة عام 2030. وهناك أربعة منظورات جديدة في التقييم تستحق اهتماماً خاصاً. والمنظور الأول هو مفهوم التغيير التحويلي. وهناك اتفاق عام على أن تحقيق أهداف التنمية المستدامة يتطلب تغييرات تحويلية تعالج الأسباب الجذرية والدوافع النظامية الكامنة وراء الفقر والاستبعاد والتلوث. وتزداد مطالبة المقيّمين بالإجابة على تساؤلات عن التأثير على المعايير والنُظم بما يتجاوز النتائج المباشرة للمشروعات. ويحتاج المقيّمون إلى فهم الطريقة التي يمكن بها تقدير التغيير على مستوى النُظم ويُقدّم الدليل الجديد بعض النصائح في هذا الصدد.

12- والمنظور الثاني هو التعقيد والتفكير النظمي. وتتشابك بعمق حلول الفقر وعدم المساواة وتغيّر المناخ والتحديات العالمية الأخرى. ويتطلب فهم هذه التفاعلات استخدام منهجيات تقييم أكثر تطوراً، بما في ذلك علم التعقيد وتحليل النُظم لتقدير الصلات والمفاضلات. وهناك حاجة إلى فهم كيفية تأثير حصائل البرامج بالعوامل الاقتصادية والسياسية والاجتماعية الثقافية والإيكولوجية وغيرها على المستويات المحليّة والوطنية والدولية. ويتيح تطبيق النهج القائمة على النظريات في التقييم أساليب قيّمة للتعامل مع التعقيد.

13- والمنظور الثالث الذي يشكل تحدياً لبرامج التنمية الريفية ومقيّمها هو معالجة الاستدامة والقدرة على الصمود في وجه تغيّر المناخ. وتُحدّد أهداف التنمية المستدامة تغيّر المناخ كعامل مضاعف للتهديدات يمكنه عرقلة التقدم في معالجة مسائل الفقر والجوع والمساواة والصحة. ويتطلب ذلك تقييم الاستدامة والقدرة على الصمود، ويشمل تقييم قدرة البلدان والمجتمعات المحليّة والأسر على التكيف مع الصدمات المناخية التي لا يمكن التنبؤ

بها، والتي يمكن أن تحدث على امتداد فترات زمنية طويلة. وتحقيقاً لهذه الغاية، يجب أن يعمل المقيّمون مع البيانات المناخية وتحليلها عند تقدير التدخلات والسياسات.

14- والمنظور الرابع هو **العدالة الاجتماعية والتقاطعية الجنسانية**. ويدخل الالتزام بعدم ترك أحد يتخلف عن الركب في صميم خطة عام 2030. وتماشياً مع إرشادات فريق الأمم المتحدة المعني بالتقييم بشأن دمج حقوق الإنسان والمساواة بين الجنسين في التقييمات، يجب أن تراعي التقييمات جميع هذه الأبعاد، وكذلك تقاطعاتها (مثل تقاطع الخصائص الجنسانية والاجتماعية والاقتصادية الأخرى، مثل الأصل الإثني، والعمر، والطبقة، ومستوى الدخل). ويتعين أن تشمل التقييمات هذه الجوانب كي تعالج الأسباب المتعددة للتمييز والإقصاء والطريقة التي تتفاعل بها. ويساعد الأخذ بمنهجيات تشاركية على جعل منظور أصحاب المصلحة الأكثر ضعفاً وتهميشاً جزءاً لا يتجزأ من التقييم.

15- وأدخلت في الدليل عدة تغييرات في **تعريف معايير التقييم** مقارنة بالإصدارات السابقة. أولاً، يُقدم الدليل معيار "الاتساق" (الذي يُشكل الآن جزءاً من المعايير الدولية، لجنة المساعدة الإنمائية التابعة لمنظمة التعاون والتنمية في الميدان الاقتصادي، 2019)، الذي سيستخدم أساساً في التقييمات القطرية والمؤسسية/المواضيعية. وتُقدّر الأنشطة غير الإقراضية (إدارة المعرفة، وبناء الشراكات، وحوار السياسات) كمجالات فرعية في إطار "الاتساق". ومن هنا، سيستخدم الصندوق ستة من معايير التقييم الدولية: الملاءمة، والاتساق، والفعالية، والكفاءة، والأثر، والاستدامة.

16- وبالإضافة إلى ذلك، سيعتمد الصندوق معايير تعالج مهمته المحددة، مثل الابتكار، وتوسيع النطاق، والمساواة بين الجنسين وتمكين المرأة، وإدارة البيئة والموارد الطبيعية، والتكيف مع تغيير المناخ، وأداء الشركاء (الصندوق والحكومة). وعلاوة على ذلك، يُصنّف الدليل المعايير بطريقة مختلفة. ويُقدّر "الابتكار" في إطار الفعالية، بينما يُقدّر "توسيع النطاق" و"إدارة البيئة والموارد الطبيعية، والتكيف مع تغيير المناخ" في إطار الاستدامة. وسيساعد ذلك على تجنب التجزؤ المفرط والتكرار وسيفضي بالتالي إلى زيادة الإيجاز في الوثائق.

17- ثانياً، يتضمن الدليل تحديثاً لتعريف معايير التقييم والأسئلة الرئيسية المتعلقة بها، وفقاً للمعايير الدولية المحدثة وبما يتماشى مع نتائج التقييم الأخيرة في الصندوق.

18- ثالثاً، يجمع الدليل حالياً بين معياري "إدارة البيئة والموارد الطبيعية" و"تغيير المناخ والتكيف معه" في معيار واحد. ويتماشى ذلك بصورة أفضل مع تعريف القدرة على الصمود في وجه تغيير المناخ الذي وضعته الهيئة الحكومية الدولية المعنية بتغيير المناخ (2018).

19- ويعرض الدليل تعاريف مشتركة وأسئلة رئيسية لمعايير التقييم الذاتي والتقييم المستقل على حد سواء. ولذلك، لن يكون ضرورياً بعد الآن إصدار اتفاق مواعمة منفصل.

20- ويتضمن الدليل مقياساً للتصنيف من ست نقاط لمعايير التقييم. ويُقدّم إرشادات عامة لتحديد التصنيفات، مع إدراك الحاجة إلى بعض المرونة للتكيف مع السمات الخاصة بكل مشروع. وستُساعد آليات استعراضات الأقران على التحكم في التباينات بين المقيّمين. وبينما سيواصل مكتب التقييم المستقل تصنيف جميع المعايير كل على حدة، لن تُصنّف إدارة الصندوق الابتكار والأثر وتوسيع النطاق وإدارة البيئة والموارد الطبيعية والتكيف مع تغيير المناخ. وعلى الرغم من أن الإدارة لن تُصنّف هذه البنود باستخدام مقياس التصنيف المؤلف من ست نقاط، فإنها سترصدها وستقيسها في برامج الفرص الاستراتيجية القطرية وعند إنجاز المشروعات وأثناء التنفيذ على النحو المبين في القسم الفرعي جيم من القسم ثالثاً من دليل التقييم. وستكون هذه القياسات متاحة لمكتب التقييم المستقل للمساهمة في التصنيفات كجزء من التحول نحو مواعمة نهج القياس بين إدارة الصندوق ومكتب التقييم المستقل.

21- وستواصل الإدارة إجراء تقديرات للأثر في 15 في المائة من حافظة المشروعات وستُقدّم إلى المجلس تقريراً عن النتائج الموحدة في نهاية دورة تجديد الموارد. وبالإضافة إلى ذلك، ستُقدّم الإدارة تقريراً عن النتائج التي

ستقاس من خلال المؤشرات الأساسية للحصائل في الأطر المنطقية للمشروعات،¹ والتي سيجري الحصول عليها أيضا من خلال منهجية استقصائية دقيقة للإسناد كما هو مبين في الفقرة 78 من القسم ثالثا جيم من الدليل.

رابعاً- الخطوات المقبلة

- 22- **وضع اللمسات الأخيرة على الجزء الثاني.** أعد مكتب التقييم المستقل وإدارة الصندوق مسودات نسخ الفصول التي ستشكل الجزء الثاني من الدليل والتي تتناول منتجات التقييم المحددة. وسيجري الانتهاء من إعداد الجزء الثاني وتطبيقه على التقييمات الجديدة في مطلع عام 2022.
- 23- **إتاحة الموقع الشبكي والأدوات التقنية التكميلية.** ستُنشر النسخة الإلكترونية من الدليل في الموقع الشبكي لمكتب التقييم المستقل في الصندوق وستكون متاحة لعامة الجمهور. وستُترجم إلى لغات العمل في الصندوق. وسيقوم مكتب التقييم المستقل وإدارة الصندوق بإعداد صفحة شبكية تكميلية تحتوي على روابط تشعبية (Hyperlinks) مؤدية إلى إرشادات تقنية أكثر تحديدا. وستشمل هذه المراجع العروض القائمة والحلقات الدراسية المتاحة على شبكة الإنترنت وكذلك الحلقات الدراسية الجديدة والمواد التي سيصدرها مكتب التقييم المستقل وإدارة الصندوق. ويعكف مكتب التقييم المستقل بصفة خاصة على إعداد سلسلة حلقات دراسية سيقدّمها أعضاء الفريق الاستشاري للتقييم. وستوفر الأدوات التقنية التكميلية لمستخدمي الدليل المهتمين مرونة في الوصول إلى المواد المتخصصة دون إقبال كاهل النص الرئيسي للدليل.
- 24- **دورة تدريبية على شبكة الإنترنت وأحداث تعليمية مركزة.** يعكف مكتب التقييم المستقل على إعداد دورة تدريبية على الإنترنت بحسب الوتيرة التي يحددها المستخدم نفسه. وستقدم الدورة التدريبية محتويات الدليل الرئيسية وستسهل الوصول إليها. وستتضمن الدورة قسما للاختبارات لتمكين المستخدمين من التحقق من مستوى الفهم والاستيعاب، والحصول على شهادة تدريب. وبالإضافة إلى ذلك، سيجري تنظيم أحداث تعليمية مشتركة بين إدارة الصندوق ومكتب التقييم المستقل لتوعية موظفي الصندوق والشركاء الإنمائيين الوطنيين بالسمات الرئيسية للدليل وأهمية التقييم في تعزيز النتائج الإنمائية.

¹ استعيض في عام 2017 بالمؤشرات الرئيسية عن نظام إدارة النتائج والأثر وإرشادات الإطار المنطقي المرتبطة به. وفي عام 2017، ومع ترحيل جميع الأطر المنطقية من النسخ الورقية/الشكل المحمول PDF إلى نظام إدارة النتائج التشغيلية، حولت جميع مؤشرات نظام إدارة النتائج والأثر إلى مؤشرات رئيسية في جميع الأطر المنطقية للمشروعات الجارية. وتمثل المؤشرات الرئيسية للحصائل مجموعة فرعية من المؤشرات الرئيسية، وستصبح إلزامية لجميع المشروعات الجديدة اعتبارا من عام 2022. وتُجمع نتائج المؤشرات الرئيسية للحصائل من خلال ثلاثة استقصاءات أثناء تنفيذ المشروع: في مرحلة خط الأساس للمشروع، ومنتصف المدة، ومرحلة الإنجاز. ووضعت منهجية خاصة بالصندوق لهذه الاستقصاءات ويمكن الرجوع إليها في الملحق الثالث من الجزء الأول من الدليل.

Main Report

Contents

Abbreviations and acronyms	2
INTRODUCTION	3
A. Why this manual?	3
B. For whom is this manual written?	3
C. What does it contain?	3
I. Setting the Foundations	6
A. The role of IFAD	6
B. Evaluation's role in supporting IFAD to achieve its development objectives	9
C. Core evaluation principles	10
II. IFAD's evaluation and learning system	12
A. IFAD's evaluation architecture	12
B. Enhancing learning and evidence use	14
III. Methodological fundamentals	19
A. Defining the evaluand and the evaluation scope	20
B. Defining the evaluation design approach	21
C. Evaluation criteria, key questions and ratings	29
D. Reviewing evaluability and data availability	36
E. Approaches and methods for data collection and analysis	37
F. The Role of Information and Communication Technologies (ICTs) for data collection and analysis	43
G. Evaluation conclusions and recommendations	46
Annexes	
I. IFAD Internal evaluation architecture	47
II. Evaluation that are responsive to social justice and gender equality	51
III. Core outcome indicators	54

Abbreviations and acronyms

CI	Core Indicators
COI	Core Outcome Indicators
COSOP	Country Strategic Opportunities Programme
CSPE	Country Strategy and Programme Evaluation
DEF	Development Effectiveness Framework
ECG	Evaluation Cooperation Group
GIZ	German Corporation for International Cooperation
ICT	Information and Communication Technology
IEG	Independent Evaluation Group
IFI	International Financial Institution
IOE	Independent Office of Evaluation of IFAD
M&E	Monitoring and Evaluation
OECD-DAC	Organisation for Economic Co-operation and Development – Development Assistance Committee
OPR	Operational Policy and Results Division of IFAD
PCR	Project Completion Report
PMU	Project Management Unit
QCA	Qualitative comparative analysis
RIA	Results and Impact Assessment Division of IFAD
RMF	Result Management Framework
SDG	Sustainable Development Goals
SKD	Strategy and Knowledge Department of IFAD
ToC	Theory of Change
UNEG	United Nations Evaluation Group

IFAD Revised Evaluation Manual – Part 1

Introduction

A. Why this manual?

1. IFAD is committed to making a significant contribution to eradicate poverty and hunger in rural areas of developing countries, while also positively impacting gender equality, climate and social justice. The main purpose of IFAD's Revised Evaluation Manual (2021) is to ensure the quality, consistency, rigor, and transparency of the evaluation function at IFAD to ultimately enhance the effectiveness of IFAD's work to contribute to the well-being of the poor in rural areas.
2. This manual is a living electronic document that will be adapted over time to reflect evolving practice, needs and lessons. IFAD's Revised Evaluation Manual represents a major revision from the 2015 edition. Changes in the latest version aim to improve implementation of IFAD's evaluation policy to which it is aligned. The manual seeks to renew, update and consolidate current guidelines. For the first time, it provides a comprehensive institution-wide approach through which self and independent evaluations are planned, conducted and used. The manual gives renewed emphasis to the importance of harmonizing and streamlining the two and maximizing the use of findings and lessons when planning and implementing projects and programmes.
3. The revision of the 2015 IFAD Evaluation Manual was undertaken by the Independent Office of Evaluation and IFAD's Management in recognition of the dynamic environment in which IFAD operates, and in response to evolution in the approaches and methodologies of international development evaluation practices. It will help ensure that IFAD's methodological practice remains state of the art.

B. For whom is this manual written?

4. The evaluation manual sets standards for self- and independent evaluations at IFAD. Its main audience is IFAD's staff and consultants, who manage evaluations or are involved in independent- and self-evaluation processes.
5. The secondary audience of this manual includes IFAD's stakeholders and partners, such as governments, private sector, civil society and development partners involved in IFAD's planning, monitoring and evaluation processes.
6. The manual may also be of interest to those who are external to IFAD and involved in the evaluation of rural development programmes. This includes Member States, international organizations, national non-governmental partners and beneficiaries and rural development practitioners.

C. What does it contain?

7. The manual presents how evaluation is performed at IFAD, and therefore it is not a sourcebook on evaluation in general. It includes essential guidance on evaluation fundamentals and criteria that are applicable to all evaluations. It also contains a section on organizational learning, recognizing that reports are of limited value if the knowledge therein is not appropriately used by as many people as possible. To this end, it provides detail on the complementarities of IFAD's self- and independent-evaluation systems and related evaluation products with a view to strengthening accountability, learning and overall utility of evaluations efforts. It also comprises specific methodological guidance on all evaluation products.
8. Readers are encouraged to read all chapters in order to get a good understanding of how evaluations should be interpreted, managed, conducted and used. The manual also serves as a reference document for information about specific issues or evaluation products.

9. The manual is divided into two parts:
- **Part 1 (Chapters I-III)** provides the overall context for evaluation in IFAD. It covers a number of foundational elements, including IFAD's mission; its evaluation objectives, architecture, frameworks, principles and criteria that guide all evaluations within IFAD.
 - **Part 2 (Chapters IV and following)** provides practical guidance on various self- and independent-evaluation products. These individual chapters can be used in sequence or as individual pieces and are intended to be living and continuously evolving documents to support evaluation in IFAD.

Part 1: Evaluation in IFAD

Part 1 introduces the foundations for evaluation in IFAD and comprises three chapters. Chapter I puts IFAD's evaluation efforts in the context of IFAD's endeavour to contribute to inclusive and sustainable rural transformation, and its contribution to the United Nations 2030 Agenda for the Sustainable Development Goals. It presents IFAD's theory of change, and introduces the purpose of evaluation and its overarching principles. [Chapter II](#) presents IFAD's institution-wide evaluation and learning systems, its different functions and types of evaluation. [Chapter III](#) explains the methodological fundamentals applicable to all evaluations.

I. Setting the Foundations

1. The 2030 Agenda unequivocally reinforces the call for increased attention, cooperation and investment in rural development (Box 1). IFAD contributes to lifting poor rural people out of poverty. No poverty eradication and inclusive growth agenda can succeed without serious attention to rural areas and sectors, which support the livelihoods of small-scale producers. Indeed, poverty has multiple dimensions that go beyond low levels of income, consumption and material assets; this is why IFAD targets its investments towards inclusive rural transformation, which is a comprehensive process during structural transformation of economies with social as well as economic implications.²

Box 1

Excerpts from the preamble of the 2030 Agenda

We are resolved to free the human race from the tyranny of poverty and want and to heal and secure our planet. We are determined to take the bold and transformative steps which are urgently needed to shift the world onto a sustainable and resilient path. As we embark on this collective journey, we pledge that no one will be left behind. The 17 Sustainable Development Goals and 169 targets which we are announcing today demonstrate the scale and ambition of this new universal Agenda. They seek to build on the Millennium Development Goals and complete what these did not achieve. They seek to realize the human rights of all and to achieve gender equality and the empowerment of all women and girls. They are integrated and indivisible and balance the three dimensions of sustainable development: the economic, social and environmental. [...]

We will devote resources to developing rural areas and sustainable agriculture and fisheries, supporting smallholder farmers, especially women farmers, herders and fishers in developing countries, particularly least developed countries.

Source: <https://www.un.org/development/desa/dspd/2030agenda-sdgs.html>

A. The role of IFAD

2. The mission of IFAD (hereafter also named Fund) is to facilitate both public and private investments, support national and global policy processes, generate and share knowledge, and develop partnerships, all in pursuit of transforming agriculture, rural economies and food systems to make them more inclusive, productive, resilient and sustainable.
3. IFAD is a specialized agency of the United Nations and an international financial institution (IFI) focused exclusively on reducing poverty and food insecurity in rural areas through agricultural and rural development. IFAD's portfolio targets small-scale producers, owners of small- and medium-sized rural businesses, and rural vulnerable groups, such as women, youth, indigenous peoples and persons with disabilities.³
4. IFAD has been ranked the top development cooperation (among 49 institutions) by the Center for Global Development in their QuODA 2021 (Quality of Official Development Assistance), based on an assessment of four dimensions of quality: prioritization; ownership; transparency and untying; and evaluation. The ranking specific to the evaluation dimension placed IFAD as the fourth development cooperation in terms of the quality of providers' learning and evaluation systems.⁴
5. Through its [Strategic Framework 2016-2025](#), IFAD is committed to pursuing three interlinked strategic objectives: (i) increase poor rural people's productive

² Rural Development Report 2016: Fostering inclusive rural transformation.

³ The main instruments for delivery are loan-funded projects and programmes, which IFAD helps governments, beneficiaries and other stakeholders to develop and implement. IFAD also has a small grant-financing window, and a new grant policy has just been developed, and reimbursable technical assistance. Moreover, IFAD is the first fund in the UN system to receive a public credit rating (AA+ by Fitch Ratings and S&P) which will allow IFAD to strengthen its resource base and catalyse private sector finance towards the achievement of the SDGs.

⁴ QuODA 2021: Aid Effectiveness Isn't Dead Yet: <https://www.cgdev.org/publication/quality-official-development-assistance>

capacities; (ii) increase poor rural people's benefits from market participation; and (iii) strengthen the environmental sustainability and climate resilience of poor rural people's economic activities. IFAD12 (2022-2024) is the final full replenishment cycle⁵ that will operate under the current Strategic Framework.⁶

6. **IFAD12** is a critical cycle for IFAD to increase its contributions to the Sustainable Development Goals (SDGs) and deliver on its core mandate of promoting sustainable rural transformation.⁷ IFAD's vision is built upon a theory of change (ToC), which is articulated in the context of IFAD12 and included in IFAD's Results Measurement Framework (RMF12).⁸
7. The pathways to impact are represented in Figure 1. The ToC is a conceptual model, not a literal representation of a linear process. Its main purpose is to provide a conceptual framework for understanding important changes that IFAD must achieve for long-term success.

Figure 1
IFAD12 Theory of change



Source: IFAD (2021). Report on IFAD12.

8. Tier 1 – SDG contribution – IFAD maintains its ambition to make significant contributions to SDG1 (no poverty) and SDG2 (zero hunger), tracked by measures of extreme poverty, and food insecurity and productivity of small scale producers, while also positively impacting the broader development goals, especially those focused on gender equality, climate and social justice.
9. Tier 2 – Developmental impact – assumes that success in contributing to global poverty reduction and food security targets, is achieved by country-level outcomes and impact. To be successful IFAD must:

⁵ IFAD's core financing is drawn from several sources. These include contributions from Member States and other donors, investment income and loan reflows every three years. Based on these financial resources, IFAD operations are planned on a three-year replenish period. IFAD12 is the 12th replenishment cycle covering the period from 2022 to 2024 (for more details about IFAD12 see: <https://www.ifad.org/en/ifad12/>).

⁶ The SDGs and IFAD Strategic Framework form the reference documents that set IFAD's longer-term ambitions.

⁷ For the 12th Replenishment of IFAD's Resources, IFAD established a target of reaching 127 million people with its operations. See Report of the Consultation on the Twelfth Replenishment of IFAD's Resources - Recovery, Rebuilding, Resilience, December 2020. This is the first time that IFAD will report against a specific theory of change. The ToC and the relevant sections in the Evaluation Manual will be revised when the new replenishment starts.

⁸ This is the first time that IFAD reports against a specific theory of change. The ToC and relevant sections in the manual will be revised when the new replenishment starts.

- **Expand impact:** increase outreach and speed up delivery to accelerate progress towards ending poverty and hunger.
 - **Deepen impact:** target the poorest and most vulnerable groups; strengthen systems and people's resilience in the face of shocks and stressors, and ensure that impact is sustainable. Environment and climate change, gender, nutrition and youth are critical and intersecting areas of work towards reducing poverty and hunger, and fostering resilient rural livelihoods. Thus, IFAD identifies four mainstreaming themes – youth, gender, nutrition and climate – as central elements to deepening impact and transforming the lives of rural populations. In practice, the mainstreaming agenda means that the ToC of projects at the design phase needs to clearly show synergies and intersectionality between different mainstreaming themes.
10. IFAD's developmental impact relates to the impact of IFAD-funded operations and is measured through independent evaluations,⁹ Management's Impact Assessments, and monitoring of outputs and outcomes within IFAD's Core Indicators Framework. Yet, the Fund is situated in a wider global development policy context that is complex, contested and non-linear. Evaluation efforts must take into account the web of actors, conflicting interests and systems interactions that enable or constrain IFAD's impact (implications for evaluation are presented in Chapter III).
 11. Tier 3 – Operational pillars – in the IFAD12 ToC highlights that, transformational country programmes are needed to drive transformative results for the rural poor. To achieve meaningful impact, country programmes must: (i) integrate inclusive approaches aimed at leaving no one behind, (ii) deepen partnerships, and enhance government ownership through a suite of adaptive products and tools suited to country needs, and (iii) significant investment in innovation and risk.
 12. **The focus on resilience to shocks and stressors as well as transformational change requires evaluators to have an operational definition of transformational results.** This is not explicitly defined in IFAD12 but can be derived from it¹⁰ and complemented with definitions provided by the Independent Evaluation Group (IEG) of the World Bank¹¹ and the German Corporation for International Cooperation (GIZ).¹² These sources are established references in international evaluation and combine social, environmental and climate change resilience aspects, and place the importance of systemic changes at the forefront.
 13. **Transformational results** are those that lead to a deep, systemic and sustainable change with the potential for large-scale impact at the national or global level. Transformative results ultimately generate changes that are profound enough to shift societies onto fundamentally different development pathways, converting a current (ecological, social, political, economic, scientific or technological) system into a fundamentally new one that forms the new mainstream.

⁹ IFAD is a results-driven organization as evidenced by the focus on results introduced in its first Development Effectiveness Framework (DEF). As the institution diversifies its instruments and enhances its ambition, it has also updated its DEF in 2021. The updated DEF will capture evolving priorities and new areas of work to ensure that the institution's approach to results is all-encompassing. IFAD's success will be assessed against the agreed indicators of the IFAD12 RMF.

¹⁰ According to IFAD12, conditions to achieve transformational results are: (i) focus on people's resilience to ensure sustainability and impact even in the face of a crisis; (ii) nurturing partnership with governments, the private sector, civil society and the non-governmental organization community, think tanks and other development organizations. The RDR 2016 also provides a definition for rural and structural transformation. <https://www.ifad.org/en/web/knowledge/-/rural-development-report-2016-fostering-inclusive-rural-transformation>

¹¹ IEG World Bank (2016) Supporting Transformational Change for Poverty Reduction and Shared Prosperity.

¹² GIZ (2019) Transformative Project Design.

14. **Proximity and adaptability** underpin the ToC and are cross-cutting principles for results delivery. Focusing on these two principles will enable IFAD to expand and deepen the results achieved when working through country programmes.
- **Proximity:** working closer to all its partners and members of marginalized and vulnerable communities to facilitate the ability to work in genuine partnership and find solutions to common problems and make a more transformational impact on policy.
 - **Adaptability:** adopting an adaptive approach to “doing development.” Adaptive management approaches emphasize the ability to proactively and reactively learn, respond to changes and evolve quickly. Rather than adopting fixed targets, as a results-based approach would, project teams and governments should be encouraged to adapt the means to achieving end goals based on lessons learned along the way.
15. Consequently, **evaluation efforts¹³ must support IFAD to implement course corrections more quickly when risks emerge that could undermine development objectives and outcomes, or when economic or other shocks take place.**¹⁴ Both principles call for more frequent planning with stakeholders and more agility during project design and implementation. This reinforces the role of evaluation, not only to assess IFAD’s impact and the role of partners, but also to navigate uncertainty and complexity of operations, to understand the pathways to impact, and to examine and assess what works, for whom, where, when and why. [Chapter III](#) presents IFAD’s evaluation approach.

B. Evaluation’s role in supporting IFAD to achieve its development objectives

16. Evaluation plays an important role in IFAD’s business model, including operational focus, corporate processes, accountability and learning systems. They are refined, adjusted and sharpened through feedback from evaluation to ensure that the Fund is in the best possible position to fulfil its mandate and meet its corporate goals.
17. IFAD’s 2021 Evaluation Policy constitutes the overall framework for evaluations within the institution¹⁵. It outlines the roles and responsibilities for evaluation and includes IFAD’s evaluation ToC. [Chapter II](#) presents and develops the key tenets of IFAD’s evaluation and learning system, including its various functions, types of evaluation undertaken and feedback loops.
18. In supporting IFAD to achieve greater development effectiveness, the Evaluation Policy identifies two primary purposes:
- Promote accountability by providing an evidence-based assessment of results achieved through IFAD lending and non-lending support and for putting in place the necessary corporate business model, policies, strategies and guidance, as well as resources and capacities to achieve these results; and
 - Contribute to enhanced learning, knowledge management and transparent feedback mechanisms to improve current and future policies, strategies, programs, projects and processes (Figure 2).
19. The 2019 report of the External Peer Review Panel on the Evaluation Function at IFAD emphasized that while accountability was a strong point of IFAD’s evaluation function, learning can be strengthened, and so can the incentives and methods through which learning loops are deployed across the institution.¹⁶ Nonetheless,

¹³ IFAD monitoring system plays a fundamental role in support of the evaluation function. This manual focuses on evaluation efforts.

¹⁴ IFAD12: Business Model and Financial Framework 2022-2024.

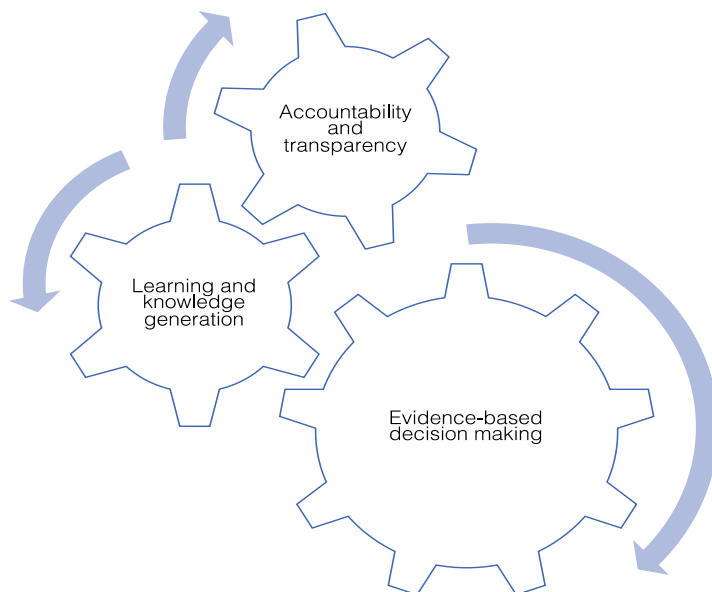
¹⁵ Revised IFAD Evaluation Policy, 2021. <https://webapps.ifad.org/members/eb/132/docs/EB-2021-132-R-5-Rev-1.pdf>

¹⁶ <https://webapps.ifad.org/members/eb/124/docs/EB-2018-124-R-8.pdf>

evaluation helps IFAD strengthen planning and implementation of its policies, strategies and operations by determining the relevance and fulfilment of its development objectives. As illustrated in **Figure 2**, both accountability and learning functions aim to promote a results-based culture, evidence-based planning and adaptive management across IFAD.

Figure 2

Accountability and learning for evidence-based decision-making



Source: This manual.

20. As further explained in **Chapter II**, IFAD's evaluation architecture comprises the independent and self-evaluation systems. The Independent Office of Evaluation (IOE) conducts independent evaluations, whereas IFAD's self-evaluation system is a responsibility of IFAD's Management. The two systems work jointly through harmonization of processes and consultations at key stages of evaluations, consistent with the independence of IOE. The range of IFAD evaluation products (see **Part 2** of this manual) ensures that both dimensions of evaluation are adequately covered.

C. Core evaluation principles

21. IFAD's Evaluation Policy (2021) identifies six interrelated key principles that underpin the organization's approach to evaluation and provide the conceptual framework within which evaluations are carried out. The principles are: **usefulness; impartiality and credibility; transparency, partnership, consultation and collaboration; evaluability; and value for money/cost effectiveness.**¹⁷
22. IFAD also subscribes to the overarching norms and standards adopted by the United Nations Evaluation Group¹⁸ (UNEG), the Evaluation Cooperation Group (ECG) of the Multilateral Development Bank,¹⁹ and the quality standards and principles of the Organisation for Economic Co-operation and Development²⁰ (OECD). Consistency and compliance with these broader principles is at the core of IFAD's evaluative work.

¹⁷ <https://webapps.ifad.org/members/eb/132/docs/EB-2021-132-R-5-Rev-1.pdf>

¹⁸ UNEG Norms and Standards for Evaluation, 2016. <http://www.unevaluation.org/document/detail/1914>

¹⁹ Evaluation Cooperation Group's Good Practice Standards for the Evaluation of Public Sector Operations, 2012. <https://ecgnet.org/content/public-sector-operations>

²⁰ OECD-DAC. <https://www.oecd.org/dac/evaluation/>

Quality standards: <https://www.oecd.org/development/evaluation/qualitystandards.pdf>
Principles: <https://www.oecd.org/development/evaluation/2755284.pdf>

23. **Evaluation ethics** is the fundamental principle underlying the six mentioned above in evaluation practice. It is the responsibility of the evaluation team to uphold ethical codes of practice, guidelines and principles. UNEG defines ethics as “**the ‘right’ or agreed principles and values that govern the behaviour of an individual within the specific culturally defined context within which an evaluation is commissioned or undertaken,**” and identifies four key principles (see Figure 3): integrity; accountability; respect; and beneficence. Systematic attention to these principles helps balance the goals of evaluations and those who drive them with the rights and interests of diverse participants and their communities. In contrast, failure to systematically consider ethics throughout the evaluation cycle can have adverse consequences for intended beneficiaries of the evaluation.

Figure 3

IFAD Evaluation function endorses UNEG’s ethical principles for evaluation

INTEGRITY: the active adherence to moral values and professional standards, which are essential for responsible evaluation practice

- Honesty and truthfulness in communication and actions.
- Professionalism based on competence, commitment, ongoing reflective practice and credible and trustworthy behaviour.
- Independence, impartiality and incorruptibility to mitigate or prevent conflicts of interest, bias or undue influence of others, which may otherwise compromise responsible and professional evaluation practice.

ACCOUNTABILITY: the obligation to be answerable for all decisions made and actions taken; to be responsible for honouring commitments, without qualification or exception; and to report potential or actual harms observed through the appropriate channels

- Transparency of the evaluation, thereby increasing accountability for performance to the public.
- Responsiveness as questions or events arise. Where corruption, fraud, sexual exploitation or abuse or other misconduct or waste of resources is identified, it must be referred to appropriate channels.
- Taking responsibility for meeting the evaluation purpose and for actions taken.
- Justifying and fairly and accurately reporting decisions, actions and intentions to stakeholders.

RESPECT: involves engaging with all stakeholders of an evaluation in a way that honours their dignity, well-being and personal agency while being responsive to their sex, gender, race, language, country of origin, LGBTQ status, age, background, religion, ethnicity and ability and to cultural, economic and physical environments

- Access to the evaluation process and products by all relevant stakeholders – with due attention to factors that can impede access such as sex, gender, race, language, country of origin, LGBTQ status, age, background, religion, ethnicity and ability.
- Meaningful engagement and fair treatment of all relevant stakeholders in the evaluation processes, so they can actively inform the evaluation approach and products rather than being solely a subject of data collection.
- Fair representation of different voices and perspectives in evaluation products.

BENEFICENCE: means striving to do good for people and planet while minimizing harms arising from evaluation as an intervention.

- Explicit and ongoing consideration of risks and benefits from evaluation processes, products and longer-term consequences.
- Maximizing benefits at systemic (including environmental), organizational and programmatic levels.
- Doing no harm and not proceeding with an evaluation when harms cannot be mitigated.
- Ensuring evaluation makes an overall positive contribution to human and natural systems and to the mission of the United Nations.

Source: United Nations Evaluation Group, 2020: Ethical Guidelines for Evaluation.

II. IFAD's evaluation and learning system

24. IFAD Evaluation Policy sets the broad framework through which evaluative evidence is produced and used. The policy emphasizes the need for effective use of and learning from evaluation products. Similarly, the use of evidence as the basis for decisions on the design and implementation of projects, programmes, and strategies is at the core of [IFAD Strategic Framework 2016-2025, the Development Effectiveness Framework](#), and **IFAD11 and IFAD12 Replenishments**. This chapter presents IFAD's evaluation architecture, its components, functions and types of evaluations undertaken. It also identifies key processes for knowledge generation and evidence use.

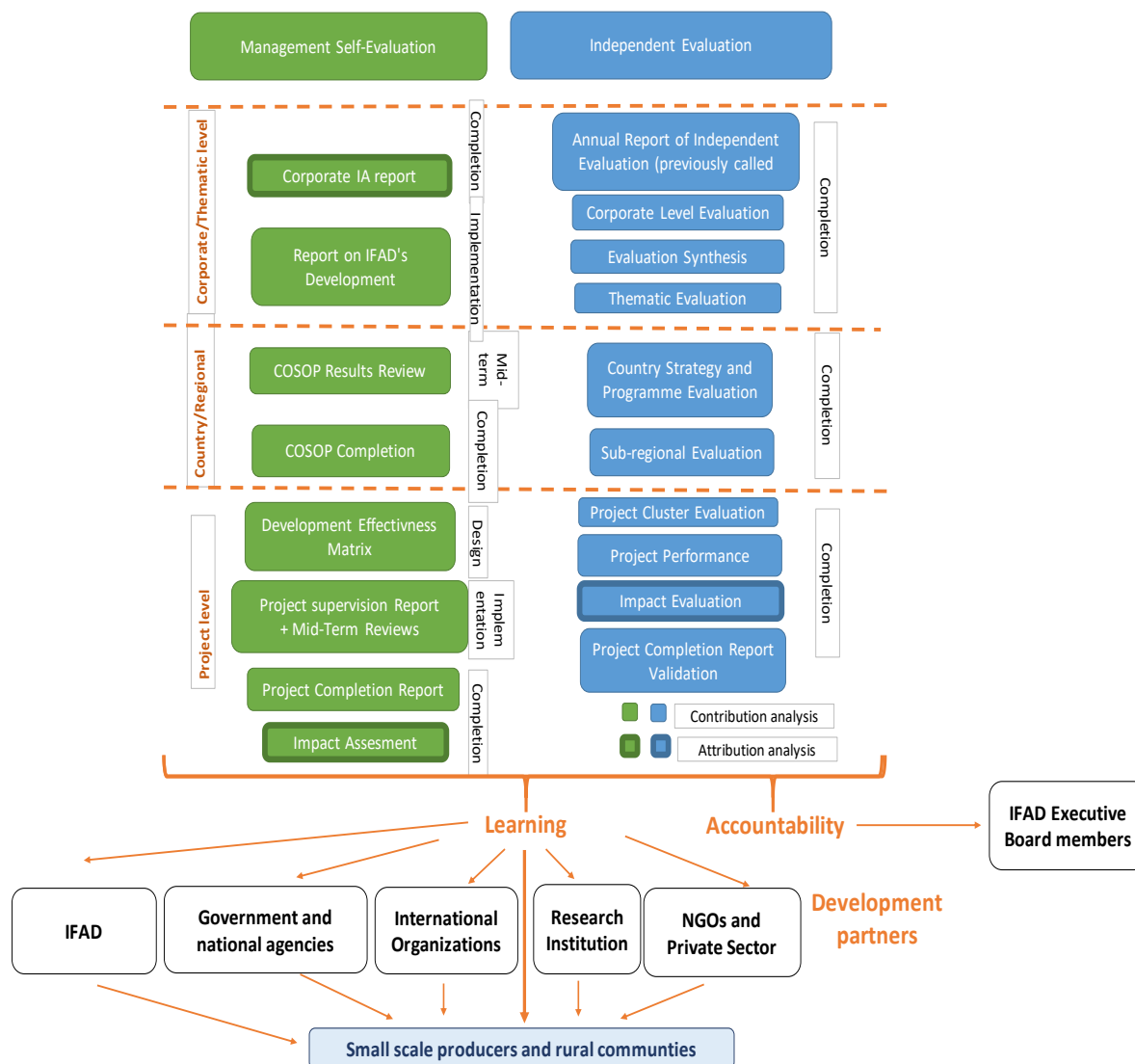
A. IFAD's evaluation architecture

25. IFAD's evaluation architecture comprises the independent and self-evaluation systems, which provide important tools for accountability, learning and knowledge management with useful practical application at strategic as well as operational levels.
26. Figure 4 shows the IFAD's evaluation architecture that combines independent and self-evaluation, as well as linkages to development partners and to IFAD's ultimate clients – small-scale rural producers and their communities.
27. **Independent Evaluations** are conducted by the Independent Office of Evaluation (IOE) which is structurally, functionally, and behaviourally independent from IFAD Management. From a governance perspective, IFAD's Evaluation Policy states that IOE reports directly to IFAD's Executive Board and that the Evaluation Committee (EC) supports the Executive Board on evaluation matters. IOE ensures that the whole evaluation function at IFAD follows internationally recognized good standards and practices. Independent evaluations help reveal what has been achieved, what does or does not work and why, and guide the development of successful policies and strategies to support rural transformation. The target audience of independent evaluations are IFAD's management and governing bodies, member countries and the international development community at large.
28. **The self-evaluation system** is a responsibility of IFAD Management and is conducted by the Operational Policy and Results Division (OPR), and the Research and Impact Assessment Division (RIA).²¹
29. Self-evaluation serves three primary purposes: (i) to obtain real-time feedback on performance and inform decision-making; (ii) to learn from experience and improve the development effectiveness of operations; and (iii) to report to IFAD's Governing Bodies on aggregated results against targets agreed upon with Members, as well as the attributable results and impact of its operations. Activities related to the first objective are carried out by project teams in regional divisions of the Program Management Department, while OPR provides guidance and support. Activities related to the second objective are led by OPR and the Strategy and Knowledge Department (SKD), and involve project teams. The third objective is the responsibility of OPR (results) and RIA (impact). Beyond providing robust results reporting for accountability, self-evaluation products are one of the sources of information for independent evaluations.
30. Strong monitoring and evaluation systems (M&E) – which promote adaptive management and learning (or Monitoring Evaluation Accountability and Learning MEAL) – are of paramount importance. In addition, the impact assessments conducted by RIA that report on attributable impacts of IFAD's investments on its goal and strategic objectives make IFAD the only IFI with this type of systematic and rigorous approach to corporate reporting. High-quality self-evaluation products

²¹ Within IFAD's structure, OPR is part of the Programme Management Department, while RIA is part of SKD.

are critical for effective IFAD-wide evaluations, as well as evidence-based programming.

Figure 4
A simplified scheme of IFAD's evaluation architecture (up to IFAD11)



Source: Elaborated by this manual.

31. Both independent and self-evaluation functions work at three different levels: (i) project; (ii) country or regional; and (iii) corporate or thematic. The details of each evaluation product are provided in Part 2 of this manual. The first level of project evaluation provides the basis of analysis for the other levels. Self-evaluation is conducted at the design, implementation and completion stages of the project. Furthermore, the majority of evaluations are based on contribution analysis that aims to measure the contribution of the interventions to the overall change. To complement these analyses, impact assessments or evaluations, conducted respectively by RIA and IOE, and corporate impact assessment reporting conducted by RIA are based on attribution analysis. They evaluate the impacts that are attributable to the interventions isolated from all factors that might have affected the outcome at both project and corporate levels. A thorough discussion on contribution and attribution analyses is presented in Chapter III.
32. The entire evaluation architecture provides evidence for accountability, learning and knowledge internally at IFAD for strategic and operational purposes, as well as externally for development partners and, eventually, for end clients. Development partners are governments and national agencies, international organizations (e.g.

other United Nations organizations), research institutions, Non-Governmental Organizations and the private sector. All evaluation products are publicly available to contribute to evidence-based programming by all partners. IFAD's end clients are small-scale producers and rural communities, which directly benefit from evidence generated by IFAD evaluation functions, through improved project design, and indirectly through improved evidence available to all development partners. The full description of the evaluation system, including its components, and the linkages and learning loops between the independent and self-evaluation systems, is contained in Annex I of part 1 of this Manual.

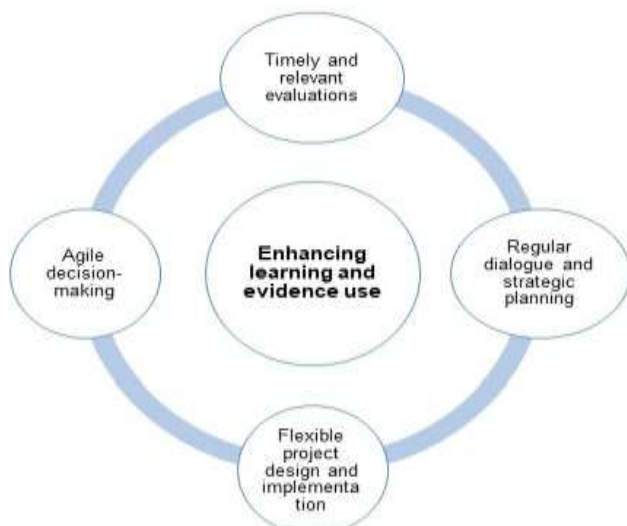
B. Enhancing learning and evidence use

33. The ultimate purpose of evaluation is to inform decision-making and contribute to broader knowledge base within IFAD, and among IFAD's external stakeholders, such as national and international development partners. A thorough and rigorous evaluation and the production of a good report are not enough for an evaluation to be useful.
34. In line with international practice,²² IFAD aims to maximize the use of evidence and evaluations throughout the entire evaluation process. Adopting an adaptive, learning-centered approach requires IFAD to learn and respond quickly and effectively to evidence and lessons. As illustrated in Figure 5, the generation and use of evidence and learning must be a continuous process to ensure IFAD becomes more agile, responsive, innovative and effective in the solutions it brings. IFAD needs to: (i) generate timely and relevant evidence; (ii) foster dialogue and strategic planning with development partners; and (iii) ensure flexible project design and implementation that constantly react to emerging results and learning, as well as external changes and events.
35. The use of evidence for quality decision-making is also at the core of the updated Development Effectiveness Framework (DEF 2.0).²³ In the DEF 2.0, all objectives and activities proposed to enhance existing self-evaluation tools have a strong focus on learning (as a cross-cutting principle) and the necessary incentives, tools, mechanisms and approaches to ensure that learning is prioritized and prized by IFAD's staff, government partners and beneficiaries. IFAD's Management is working to ensure that a culture of learning is strengthened in the coming years.

²² See for example: World Bank (2019) World Bank Group Evaluation Principles.

²³ <https://webapps.ifad.org/members/ec/115/docs/EC-2021-115-W-P-6.pdf>

Figure 5
Learning and evidence use — an iterative process



Source: Elaborated by this manual.

36. The organizational and functional independence of IOE is essential to ensure credible, solid and transparent evidence in line with international norms and standards and the principles set out in IFAD's Evaluation Policy. At the same time, leveraging all components of the evaluation system can generate a virtuous learning circle for IFAD. As such, the collaboration between IOE and IFAD Management is key to ensuring the relevance, timeliness and utility of the evaluations and boost that virtuous learning loop. In this light, the pathways and processes outlined below do not hamper the independence of IOE and aim to introduce innovative elements in the evaluation processes and learning loops, which complement the existing ones.
37. Drawing from and contributing to IFAD's Knowledge Management Strategy (Box 2), synergies and complementarities between independent and self-evaluations are maximized in two broad action pathways:
- evidence and knowledge generation; and
 - knowledge use within and beyond IFAD.
38. Together, these pathways aim to lead to the creation of an evidence base that is useful to both IFAD and its partners and is systematically embedded in IFAD's operations.

Box 2

IFAD Knowledge Management Strategy (2019)

IFAD's Knowledge Management Strategy for the period 2019-2025 is part of the organization's approach to increase its development effectiveness. The strategy and its accompanying three-year action plan will help IFAD to leverage the best available and most relevant knowledge, based on both evidence and practice, from its own work, with partners and from other external sources.

Activities will be implemented in three broad areas: knowledge generation, knowledge use, and building the enabling institutional environment for evidence-based learning and knowledge sharing.

Source: <https://www.ifad.org/en/-/document/knowledge-management-strategy>

39. Different evaluations require different levels of engagement at different points in time. In addition, the implementation of specific processes and activities may differ according to the type of evaluation and the needs and contexts of specific institutional operating and client environments, as discussed in Part 2 of this

manual. Yet, there are overarching approaches that set the basis for knowledge generation and use, as described in Pathways 1 and 2.

Pathway 1: Evidence and knowledge generation – IOE-MANAGEMENT collaboration

40. To maximize the synergies between independent and self-evaluations, collaboration between IOE and IFAD Management (Management) is pursued throughout the evaluations process, consistent with the independence of IOE. This includes the processes for selecting, planning and designing evaluation products, conducting evaluations and ensuring identification and sharing of lessons learned and recommendations.

Box 3

Evaluation Selection, Planning and Design
<ul style="list-style-type: none"> ■ IOE and Management cooperate to identify evaluation needs and demands. Independent evaluations and thematic oriented self-evaluations will be planned based on the strategic and operational needs of IFAD and external stakeholders. For example, IOE and Management collaborate to identify evaluation topics. ■ IOE and Management collaborate to prepare multi-year strategy and annual evaluation plans (self and independent) for Executive Board approval, and to improve relevance and timing of different evaluation products, for example, between Country Strategic Opportunities Programmes (COSOPs), CCRs and country strategy and programme evaluations (CSPEs).
Conducting Evaluations
<ul style="list-style-type: none"> ■ Synergies and collaboration are pursued throughout the conduct of evaluations. This includes exchanging information during mid-term reviews, preparing other self-evaluation products and reducing overlaps that are not beneficial.
Evaluation Recommendations and Completion
<ul style="list-style-type: none"> ■ Joint technical reviews and learning events are undertaken before recommendations are finalized. The purpose of these meetings is to draw out and discuss lessons and recommendations to ensure full internalization and learning. ■ IOE and Management jointly organize round-table workshops and/or learning events to discuss the results and lessons emerging from the relevant evaluation with multiple partners. Similarly, learning workshop are organized for corporate-level evaluation at an appropriate stage in the process.

Pathway 2: Evidence and knowledge use within and across IFAD

41. The evaluation process does not end with the evaluation report. The follow-up steps are critical for ensuring knowledge use within and beyond IFAD and making certain that findings and lessons from each evaluation are communicated, absorbed and applied across the institution and shared with development partners and end-clients in the rural areas. To this end, key activities include:

Box 4

- **Synthesis of overarching findings:** independent and self-evaluations are used to prepare syntheses of evaluations (including policy briefs and infographics) to inform relevant corporate policies, strategies and operational processes in IFAD.
- **Learning events:** IOE and Management organize capitalization workshops, both internal stock-taking events and events in collaboration with key development partners, as required, to discuss the findings of key evaluations in order to facilitate the learning and uptake of lessons.
- **Leveraging technology to learn from evaluation:** in order to enhance utilization, evidence from both self- and independent evaluations are easily accessible and in appropriate formats for the target audience. Going forward, increasing the use of information and communication technologies (ICTs) can help systematize and extract lessons from all types of independent and self-evaluations to design more impactful development interventions.
- **Management Action Register:** a formal process where IOE and Management discuss and identify actions to be implemented as a result of key recommendations from each major evaluation. Progress on these actions will be followed by after-action reviews. Management plans to bring Volume II of the President's Report on the Implementation Status of Evaluation Recommendations and Management Actions online, to create an online evaluation tracking system to allow registration of recommendations, identification of follow-up actions by management, and progress reporting. The system would also allow data extraction for quick generation of reports on the status of recommendations. Similar electronic systems are in place in other multilateral organizations (e.g. UNDP, World Bank Group).
- **After-action reviews** of recommendations and their implementation for monitoring both performance and the results of the actions taken to address recommendations. After-action reviews may involve development partners as well.
- **Operational Strategy and Policy Guidance Committee – Project and COSOP designs:** RIA's Impact Assessments create knowledge based on attributable impact analyses that feed into project and COSOP designs to improve programming.

42. In conclusion, given the emphasis of the 2021 Evaluation Policy on the collaboration between IOE and IFAD Management, Section 2.1 of this manual provides guidance on their interactions. Box 5 provides the key phases of an independent evaluation process and its interactions with IFAD Management. More details on the self- and independent evaluation steps and their interactions are presented in Part 2 of the manual. It is important to underline that other development partners and clients play a fundamental role in IFAD-funded operations and are key stakeholders in an evaluation process.

Box 5

Overview of the phases of an independent evaluation process and its interactions with IFAD Management

<p>Design</p> <ul style="list-style-type: none"> • Draft approach paper. IOE prepares an approach paper which outlines the objectives, scope, methodology and process of the evaluation. IOE interacts with the key partners, notably IFAD Management and the government (when applicable) and seeks their comments. For more complex evaluations, an inception workshop may be held and a preparatory mission may be conducted. • Final approach paper. IOE finalizes the approach paper by addressing the comments of the stakeholders and shares the paper with them.
<p>Conduct</p> <ul style="list-style-type: none"> • Field missions for data collection. IOE plans the evaluation field missions in liaison with IFAD Management, the government and other stakeholders, as required, and communicates with them in advance. • Wrap-up meetings. When country visits are conducted, at the conclusion of the visits IOE organizes a debriefing with the government, IFAD operational staff and other stakeholders, as required.
<p>Reporting</p> <ul style="list-style-type: none"> • Draft report. IOE prepares the draft evaluation report and shares it with IFAD, the government and other stakeholders, as required, for comments. • Findings and recommendations. To the extent possible, evaluation findings and recommendations are discussed internally and with stakeholders to enhance ownership and use. The purpose of these meetings is to draw out and discuss findings and recommendations in order to ensure full internalization and learning. • Management Response. IFAD Management prepares a response to the evaluation, which is included in the final report and discussed at relevant sessions of the Governing Bodies along with IOE's comments. • Final report. IOE finalizes the evaluation report by independently addressing the comments by IFAD and the government (and other stakeholders, as appropriate). The final report is shared with stakeholders together with an audit trail summarizing how the comments were addressed.
<p>Completion and dissemination</p> <ul style="list-style-type: none"> • Final workshops. IOE organizes final workshops and learning events in collaboration with IFAD Management, the government and other stakeholders (when applicable) to share and discuss the findings, lessons and recommendations. • Agreement at Completion Point. For country-level evaluations, the Agreement at Completion Point contains a summary of the evaluation findings and recommendations that IFAD Management and the government agree to adopt and implement within specific time frames. IFAD Programme Management Divisions responsible for implementing the process. IOE takes note on the progress and the final outcome. • Presentation to IFAD Governing Bodies. Selected evaluations are presented to relevant sessions of the Evaluation Committee and Executive Board. • Dissemination. IOE ensures appropriate dissemination of the evaluations through the IOE website and newsletter. When required, IOE also collaborates with the core partners to disseminate the messages and evaluative innovations and methods.

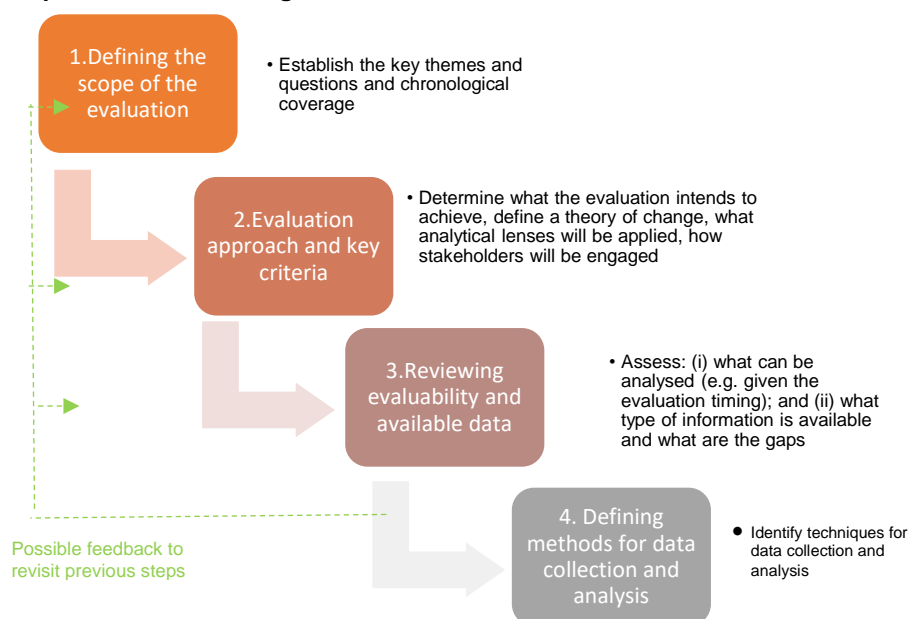
43. IFAD evaluations play an important role in IFAD's knowledge management systems as they generate and globally share knowledge on investing in sustainable and inclusive rural transformation. This, in turn, enables IFAD to play a greater advocacy role in supporting global efforts to achieve the SDGs.
44. Communication and dissemination of evidence and lessons, beyond IFAD, is therefore an integral part of the evaluation process. IFAD's evaluation policy clearly states that all independent and self-evaluation products shall be disclosed to the public and disseminated widely. At design stage, all evaluation products include a communication and dissemination plan to ensure that evaluations are shared effectively. Specific dissemination approaches for each type of self- and independent evaluation product are described in Part 2 of this manual.

III. Methodological fundamentals

45. Chapter III is dedicated to the methodological fundamentals for conducting evaluations at IFAD, ranging from scoping, design approaches, evaluation criteria and ratings used to assess different types of performance, as well as methods for data collection and analysis. It draws from contemporary literature on evaluation and from the experience of IOE and RIA. Impact assessments undertaken by RIA, which complement other evaluations with a unique approach, are based on theoretical and applied economics literatures, and are briefly introduced in a separate sub-section. The chapter concludes with a section on tips for preparing conclusions and recommendations for evaluation reports.
46. While some sections (e.g. scoping and design approaches) apply primarily to independent evaluations, they also set a reference for other IFAD divisions to be used in line with the type and scope of evaluation conducted. The sections on the evaluation criteria and ratings apply to the entire evaluation system. Awareness of the methodological fundamentals by evaluators helps set standards and ensure consistency in methodology and in reporting formats across evaluators and evaluations. It enhances the robustness and rigour of IFAD evaluation products. This chapter is not intended as a comprehensive guide to evaluation methodology and provides references to already existing methodological guidance and studies. Specific issues and considerations for different types of evaluation products are covered in Part 2 of this manual.
47. **Key steps in evaluation design.** The key steps presented in Figure 6 can apply to most evaluation exercises. They include: (i) the definition of the scope (coverage) of the evaluation (topics, time period); (ii) the establishment of an evaluation approach, the elaboration of a theory of change (ToC) and evaluation criteria; (iii) the evaluability assessment (defining what can be evaluated and what data and information are already available); and (iv) the determination of specific methods for data collection and analysis. Figure 6 also shows possible feedback loops in the design steps – the evaluation steps can be considered as iterative rather than strictly sequential. This chapter further elaborates on these steps.

Figure 6

Key steps in evaluation design



Source: Elaborated by this manual.

A. Defining the evaluand and the evaluation scope

48. While some evaluations (e.g. project-level or country-level) can be more easily standardized than others, all evaluations ought to start by defining the *evaluand* (what is being evaluated: for a thematic or strategic evaluation this may require an operational definition) and the scope of the evaluation (e.g. the specific topics to be analysed, the time interval to be considered). This is particularly important for strategic evaluations and country-level evaluations, where, for example, the evaluators will need to determine the number of projects to be reviewed and the time frame. Defining the scope and coverage may also be useful for project-level evaluations, where the analysis may need to focus on certain project components or specific topics.
49. As an example, the 2020 Corporate-level evaluation on IFAD's support to innovations for inclusive and sustainable smallholder agriculture elaborated an operational definition of innovation: "A new way of acting – practice, approach/method, process, product, or rule – brought in or implemented for the first time, considering the context, time frame and stakeholders, with the purpose of improving performance and/or addressing challenge(s)."²⁴ It also set overarching questions such as:
- To what extent (how and why) have corporate instruments, tools and approaches been successful in promoting agricultural innovations within country programmes?
 - To what extent (how and why) have IFAD's operations promoted agricultural innovations that: (a) have responded to smallholder farmers' needs/demand; and/or (b) were targeted and inclusive?
 - How have those innovations led to positive outcomes, and how have they been scaled up for sustainable and resilient development of smallholder agriculture?
50. Other examples of operational definitions that helped define the scope of strategic evaluations are presented in Box 6.

Box 6

Operational definitions in corporate evaluations and evaluation syntheses

Corporate-level Evaluation of IFAD's Financial Architecture - 2018

"Financial architecture" is defined as the policies and systems adopted to mobilize, manage, allocate and disburse financial resources to fulfil IFAD's mandate of helping to reduce rural poverty. The key elements of financial architecture can be summarized under four broad headings: (i) sources of funds; (ii) financial support instruments; (iii) allocation system; and (iv) management, oversight and governance. The financial architecture of IFAD is not a discrete "programme" or a "policy" underpinned by a dedicated logical framework. It is, rather, the result of the stratification of a number of policies and decisions made by the governing bodies and IFAD Management over forty years.

Evaluation Synthesis on Community-driven Development in IFAD-funded projects - 2019
Community-driven development is a way to design and implement development policy and projects that facilitates access to social, human and physical capital assets for the rural poor by creating the conditions for:

- transforming rural development agents from top-down planners into client-oriented service providers;
- empowering rural communities to take initiative for their own socio-economic development (i.e. building on community assets);
- enabling community-level organizations – especially those of the rural poor – to play a role in designing and implementing policies and programmes that affect their livelihoods; and
- enhancing the impact of public expenditure on the local economy at the community level.

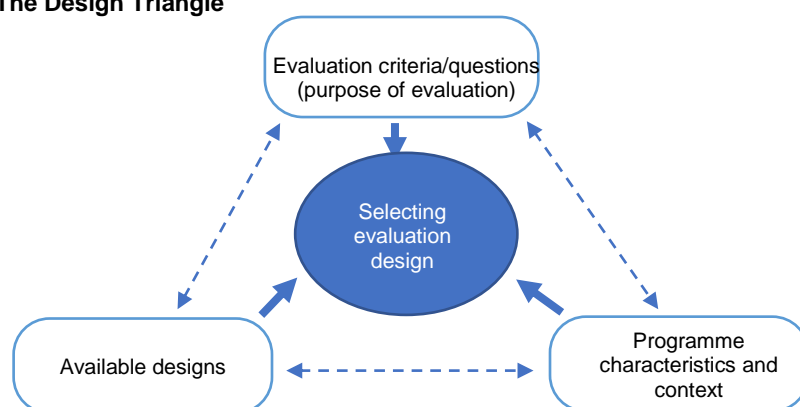
Source: excerpted from: <https://www.ifad.org/en/web/ioe/-/ifad-s-financial-architecture>; and <https://www.ifad.org/en/web/ioe/-/community-driven-development-in-ifad-supported-projects>

²⁴ <https://www.ifad.org/en/web/ioe/-/corporate-level-evaluation-on-ifad-s-support-to-innovations-for-inclusive-and-sustainable-smallholder-agricultu-1>

B. Defining the evaluation design approach

51. An evaluation approach comprises the intended objectives of an evaluation and the analytical and axiological (i.e. the system of values) lenses to be applied. Different evaluation approaches have comparative advantages in addressing specific concerns and needs. Selecting the most appropriate evaluation approach is a vital stage in the overall evaluation process.
52. The evaluation design shall match the purpose of evaluation, the evaluation questions and the nature of the intervention. Figure 7 presents the Stern et al. (2012) "Design Triangle"²⁵ pinpointing the three factors that need to be taken into account when deciding suitable evaluation designs: the evaluation questions that need to be answered; the "characteristics" of the intervention to be evaluated; and the range of available designs. For example, is the programme implemented in different settings, at different levels? A number of key considerations for evaluation design are specific to multiple-level and multi-site evaluations – e.g. country, regional, cluster evaluations – as this may involve analysing data at multiple levels of decision-making, sectors and locations. The methodology selected will enable the evaluation questions to be answered using credible evidence.²⁶ Throughout the following sub-sections, different methods options are presented to address different evaluation questions.

Figure 7
The Design Triangle



Source: Stern et al. (2012), op.cit.

53. The following sub-sections present some recent approaches and practices that have emerged and that evaluators may need to consider when defining the evaluation overarching approach, as required by the topic and context.

Emerging evaluation practices

54. Since the launch of the 2030 Agenda, there have been advancements in evaluation practices that are relevant to IFAD. **Box 7** provides a brief summary of four practices that stem from the 2030 Agenda and the Sustainable Development Goals. The first two advancements, "transformative change," and "complexity and systems thinking," are of particular importance when evaluating large-scale or global phenomena, processes and systems.

²⁵ Stern, E., et al. *Broadening the range of designs and methods for impact evaluations. Report of a study commissioned by the Department for International Development.* Working Paper 38. (London: Department for International Development, 2021).

²⁶ The Centre for Evaluating Complexity across the Nexus (CECAN) has recently launched a tool to support the process of methodological selection. See Befani, B (2020) *Choosing appropriate evaluation methods.* It also provides an overview of the potential and weaknesses of 15 methods. A tool for assessment and selection. CECAN.

55. The third advancement relates to how evaluations address the “sustainability of interventions in the context of climate change” as this has far-reaching consequences for the environment and society.
56. The fourth advancement builds on the “no one left behind” principle of the 2030 Agenda. This requires evaluation efforts to integrate social justice, gender intersectionality and culturally responsive principles, which aligns well with IFAD’s value of equity and focus on the poorest people and communities. Annex II provides additional details on social justice and gender responsive evaluations at IFAD.

Box 7

SDGs and their implications for evaluation practice²⁷

Transformative change: the 2030 Agenda recognizes the need to transform societies through sustainable, resilient and inclusive paths. There is general agreement that the achievement of the SDGs requires transformational changes at scale that address root causes, and systemic drivers of poverty, exclusion and pollution. Evaluators are increasingly called upon to answer questions about effects on norms and systems. In order to understand how interventions, programmes and policies support lasting system change, evaluators need to engage with strategic and aggregate-level evaluations and understand how systems-level change can be evaluated.

Complexity and systems thinking: the solutions to poverty, inequality, climate change and other global challenges are deeply interrelated in complex ways. For example IFAD’s ambitions to contribute to SDG2 couples natural processes with social and economic processes. Understanding these interactions require the use of more sophisticated evaluation methodologies that include complexity science and systems analysis to assess the interconnectedness and trade-offs. The focus on complexity also stresses the need to focus on the context and analyse how programme outcomes are influenced by the economic, political, socio-cultural, ecological and other factors in the local, national and international context. A paradigm shift is therefore emerging, moving from the linear, hierarchical and static logframe to a more complex and dynamic approach to examine whether and how outcomes and impact are achieved. The application of theory-based approaches to evaluation offers valuable methods for design, data collection and analysis of findings.

Sustainability and Climate Resilience: the SDGs identify climate change not only as one of its specific objectives but also as a threat multiplier with the potential to worsen progress on poverty, hunger, equality and health. Evaluating sustainability and resilience requires a different methodological approach not the conventional evaluations of programme outputs and outcomes. Both sustainability and resilience involve assessing the ability of communities or other entities to respond to, cope with and adapt to shocks and stresses, which may occur over long periods and are usually unpredictable. Evaluation efforts are seen as essential for providing evidence on whether actions to address the complex challenges related to climate change are on track for achievement of the SDGs. To this end, evaluators must engage with climatic data and analysis in the evaluation of interventions and policies.

Social justice and gender intersectionality – No one left behind: The commitment to leave no one behind is at the heart of the 2030 Agenda. Gender equality, reducing inequalities and ensuring “no one is left behind” are considered as distinct but linked core principles and goals of the SDGs. In line with the UNEG guidance on Integrating Human

²⁷ The information in the box has been sourced from several sources:

IIED, *Five considerations for national evaluation agendas informed by the SDGs* (London: IIED, 2016).

Rob D. van den Berg, Indran Naidoo, and Susan D. Tamondong, eds., *Evaluation for Agenda 2030: Providing Evidence on Progress and Sustainability* (Exeter, UK: IDEAS, 2017).

J.I. Uitto, J. Puri, and R.D. van den Berg, *Evaluating Climate Change Action for Sustainable Development: Introduction* (2017). In: Uitto J., Puri J., van den Berg R., eds. (2017) *Evaluating Climate Change Action for Sustainable Development*. Springer, Cham.

Michael Bamberger, Marco Segone and Florencia Tateossian *Evaluating the Sustainable Development Goals With a “No one left behind” lens through equity-focused and gender-responsive evaluations*. (New York: UN Women, 2016).

Rob D. van den Berg, Cristina Magro, and Silvia Salinas Mulder (eds), *Evaluation for Transformational Change: Opportunities and Challenges for the Sustainable Development Goals*. (Exeter, UK: IDEAS, 2019). See also as a reference: Patton, M.Q., *Blue Marble Evaluation: Premises and Principles*, (New York: Guilford Publishing, 2019).

Rights and Gender Equality in Evaluation and Practice,²⁸ evaluations must take into consideration all these dimensions, as well as their intersections (for example, intersection of gender and other socio-economic characteristics, such as ethnicity, age, caste and income level).

UN Women defines intersectionality as: “overlapping, concurrent forms of oppression which point to the depths of inequalities and the relationships among them in any given context. Using an intersectional lens also means recognizing the historical contexts surrounding an issue.”²⁹ Intersectionality is about:

- Fighting discrimination, ***within*** discrimination;
- Tackling inequalities ***within*** inequalities;
- Protecting minorities ***within*** minorities.

Evaluations must also incorporate these principles throughout the process in order to address multiple causes of discrimination and exclusion (e.g. age, race and ethnicity, social status, disability) and the way they interact in a specific context. Participatory methodologies, unpacking the assumptions of the power relations in evaluations, offer approaches to conduct more inclusive evaluations.^{30,31} By using participatory methodologies, evaluators will work in ways that fully consider differential impacts by gender and the way they interact with other forms of discrimination.^{32, 33}

Applying theory-based approaches

57. In line with current international practice, IFAD evaluations follow a theory-based approach.³⁴ While different definitions of theory-based evaluations exist, they are based on an explicit theory of change (ToC), which explains the theory of a development intervention. The evaluation is then designed to test the theory.³⁵ They contrast, therefore, with evaluation approaches that look solely at outcomes. Furthermore, theory-based based evaluation is part of an approach to evaluation and not a specific method or technique. It is a way of structuring and undertaking analysis in an evaluation, which helps establish whether the linkages between interventions and intended impacts are plausible, account for other contributory factors, and capture unintended effects.
58. Theory-based approaches seek to: (i) identify and explain the influence of context on program results; (ii) understand the underlying operating mechanisms that generate the observed effects and how these effects vary between different contexts and populations. As such, theory-based approaches move beyond assessing “what has changed” to also answer the more difficult questions how, why, where and for whom.

²⁸ The UNEG Handbook provides a tool to integrate human rights and gender equality into the practice of evaluation. United Nations Evaluation Group, *Integrating Human Rights and Gender Equality in Evaluations* (2014), <http://www.unevaluation.org/document/detail/1616>

²⁹ UNEG, *Integrating Human Rights and Gender Equality in Evaluation and Practice: Towards UNEG Guidance* (New York: UNEG, 2011).

³⁰ A. Stephens, E.D. Lewis and S.M. Reddy, *Inclusive Systemic Evaluation (ISE4GEMs): A New Approach for the SDG Era* (New York: UN Women, 2018).

³¹ Evaluation Cooperation Group, “Gender. Main messages and findings from the ECG Gender practitioner’s workshops” (Washington, D.C., 2017).

³² A. Stephens, E.D. Lewis and S.M. Reddy, *Inclusive Systemic Evaluation (ISE4GEMs): A New Approach for the SDG Era* (New York: UN Women, 2018) <https://www.unwomen.org/en/news/events/2018/10/ise4gems-launch-event>

³³ Evaluation Cooperation Group, “Gender. Main messages and findings from the ECG Gender practitioners’ workshops” (Washington, D.C., 2017), <https://www.ecgnet.org/document/main-messages-and-findings-ieg-gender-practitioners-workshop>

³⁴ Evaluation Cooperation Group, *ECG Big Book on Good Practice Standards*. (Washington D.C., 2012).

³⁵ Policy interventions (programs and projects) rely on underlying theories regarding how they are intended to work and contribute to processes of change. These theories usually called theories of change are often made explicit in documents but sometimes exist only in the minds of stakeholders. Program theories (whether explicit or tacit) guide the design and implementation of policy interventions and also constitute an important basis for evaluation. See for example: Jos Vaessen, Sebastian Lemire, and Barbara Befani. *Evaluation of International Development Interventions: An Overview of Approaches and Methods*. (Washington, D.C.: Independent Evaluation Group, World Bank, 2020).

59. Importantly, theory-based approaches highlight the **assumptions**, the conditions and risks that sometimes are left as implicit in the design of a project, strategy and policy. In doing so, they help identify possible gaps in the logical chain.
60. In some cases, a project, country, or corporate strategy and policy are elaborated using a ToC. When a ToC has not been elaborated explicitly, the evaluation teams may **reconstruct**, making explicit that they will have to elaborate one. In such cases, it is important to seek feedback from the main stakeholders to ensure that the reconstruction is realistic and reflects stakeholders' understanding.
61. Two important practical tips for developing a ToC are: (i) it should be relatively simple to understand and represent graphically: if it cannot be explained, it can hardly be understood, and, therefore, will not be used; (ii) it is a living instrument that needs to be revisited during the evaluation process – benefiting from stakeholders' insights.
62. Examples of ToCs are presented in Figure 8 (for evaluation synthesis prepared by IOE) and Figure 9 (for the impact assessment of a project by RIA).

Figure 8

Theory of change of the 2017 Evaluation Synthesis on Gender Equality and Women's Empowerment

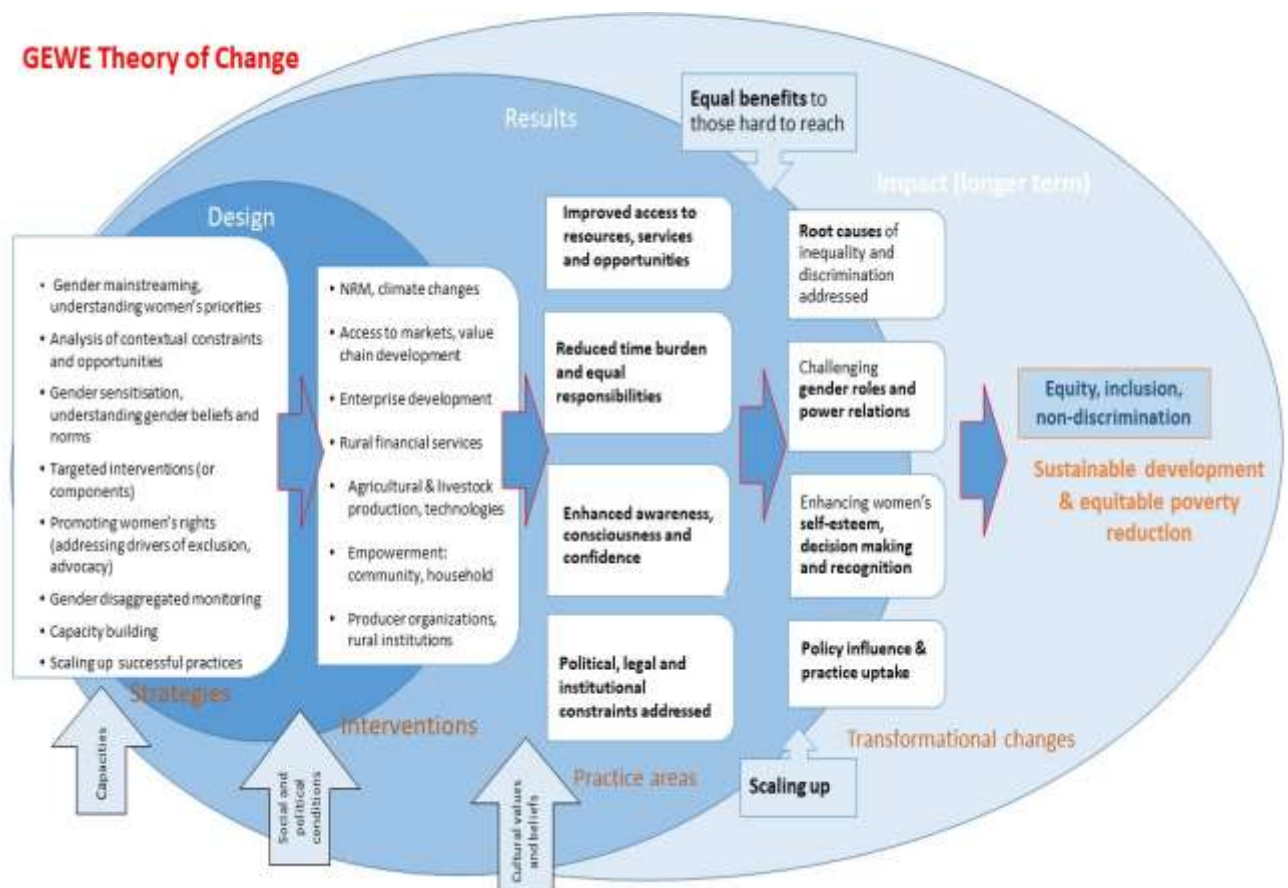
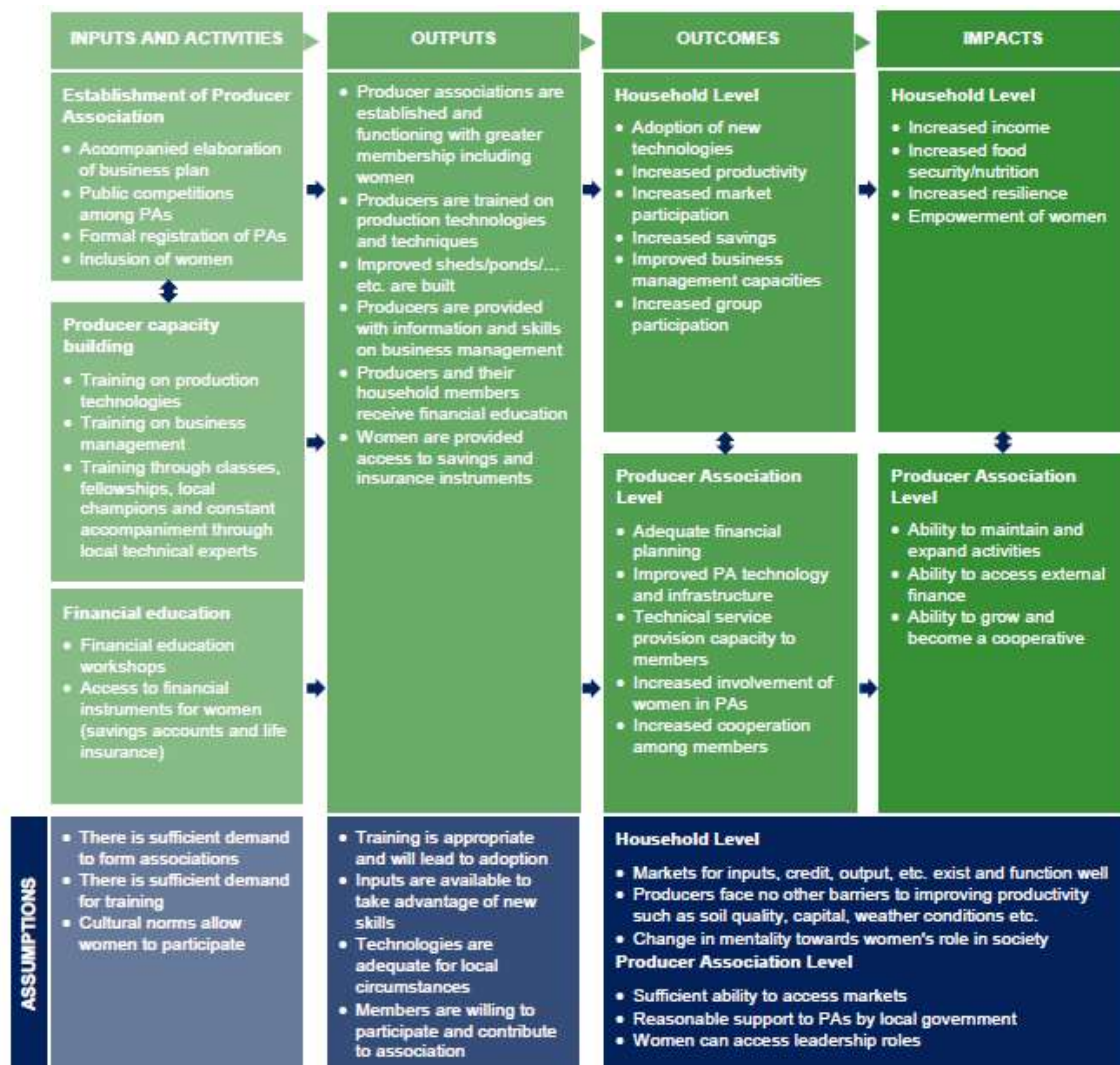


Figure 9
Theory of change from the Impact Assessment of Peru Strengthening Local Development in the Highland and High Rainforest Areas Project



63. Box 5 refers to two theory-based paradigms that are often used or referred to in literature and help operationalize theory of changes. Box 5 captures two prominent categories: realist evaluation and contribution analysis.³⁶ Recently, both perspectives were featured in guidelines by the Evaluation Cooperation Group for gender-responsive evaluations.^{37,38} These perspectives can be used in combination with several data collection³⁹ and data collection techniques (Section 3.4). For example, a case-based method can be incorporated with a theory-based design, assessing a number of different interventions, each as a separate case, and using a method such as contribution analysis (see Box 8) to assess causality for each one⁴⁰

³⁶ For a detailed practical example of how to implement and maximize the value of contribution analysis see, for example: CDI *Contribution Analysis and Estimating the Size of Effects: Can We Reconcile the Possible with the Impossible?* Practice Paper (East Sussex: Centre for Development Impact, 2019).

³⁷ Evaluation Cooperation Group, *Integrating gender into project-level evaluation* (ECG reference document (Washington D.C., 2017).

³⁸ Note that these are examples of theory-based applications. There are other theory-based evaluation approaches, for instance, in impact evaluations, for example, the 3ie portal [here](#)

³⁹ An illustrative practical example of how combined methods and designs can be found in IFAD's pilot Participatory Impact Assessment and Learning Approach evaluation (PIALA) which combined a theory-based, mixed-methods approach to evaluation that was essentially participatory.

⁴⁰ For detailed step-by-step guidelines on theory-based evaluations see, for example: *BEAM Exchange Guidelines for evaluating the impact of market systems* (the BEAM Exchange is a platform for knowledge exchange and learning about the role of markets in poverty reduction).

Box 8

Theory-based evaluation paradigms: two examples⁴¹

Realist evaluation: primarily designed to answer the question of: what works, for whom, in what respects, to what extent, in what contexts, and how? The basic message of realist evaluation is that evaluation needs to focus on understanding what works better for whom, under what circumstances; and what aspect within a programme makes it work. Under a realist perspective, an evaluator would typically investigate the underlying mechanisms that generate the effects, how they interact with the context and the main stakeholders, leading to differentiated results. They are most appropriate for evaluating new initiatives that seem to work but “where, how and for whom” is not yet understood; and/or for programmes that will be scaled up, to understand how to adapt the intervention to new contexts.

Examples:

- ITAD’s FoodTrade East and Southern Africa Regional programme Final Evaluation (2019). This evaluation combines realist approaches enquiry with contribution analysis, case studies, and thematic synthesis.
- Investment Climate Fund (2020). Portfolio Evaluation. Support for policy Change. This evaluation used a realist approach using qualitative data analysis software.

Contribution Analysis (CA): Introduced by John Mayne (see Mayne 2012), this is primarily designed to answer the question of: how far did the programme contribute to change? CA is a theory-based approach to analyse causality and it is used alongside a ToC that explicitly set out how change is, or was, supposed to happen. It is essentially a narrative approach that can be supported by various types of evidence, where the evaluator formulates and then tests a contribution story that explains how the intervention has supposedly achieved (or is supposed to achieve) its impact. The contribution story is usually visualized as a causal chain of intermediate steps or outcomes, with assumptions and risks that make each step more or less likely to materialize. CA is particularly useful in fields of work such as research, policy influencing, markets, capacity development and mobilization, where there are often many different contributors to change. A distinguishing feature of contribution analysis is the emphasis on identifying plausible alternative explanations to results observed. These could include, for example, other related government programmes, economic or social trends.

Examples:

- Contribution analysis of a Bolivian innovation grant fund: mixing methods to verify relevance, efficiency and effectiveness (Giel Ton 2017).

Systems mapping as a tool to support theory-based evaluations

64. The use of ToC is important in most evaluations but a number of key considerations for evaluation design are specific to IFAD strategic and aggregate-level evaluations, such as country-level, thematic, project-cluster and corporate evaluations. These have the following characteristics: (i) they are multi-project evaluations that often cover multiple levels of interventions, multiple sites (communities, provinces, countries) and multiple stakeholder groups at different levels and sites; (ii) they contain both summative elements as well as some formative focus and may contain important lessons for oversight bodies, management, operations or other stakeholders.⁴²

⁴¹ HM Treasury, *Magenta Book Analytical methods for use within in evaluation*. (United Kingdom, 2020)
[J. Mayne, \(2012\) 'Contribution analysis: Coming of age?', Evaluation, 18\(3\), pp. 270–280](#)

For a discussion of the practical implications of realist and contribution analysis approaches for large, multi-country complex interventions, see: Centre for Development Impact (2020) Reality Bites: Making Realist Evaluation Useful in the Real World, CDI Practice Paper 22, and (2019) *Contribution Analysis and Estimating the Size of Effects: Can We Reconcile the Possible with the Impossible?*, CDI Practice Paper 20.

⁴² Hugh Waddington, Edoardo Masset, Emmanuel Jimenez, “What have we learned after ten years of systematic reviews in international development?,” *Journal of Development Effectiveness*, (2018): 10:1, 1-16, DOI: [10.1080/19439342.2018.1441166](#)

65. For this type of evaluation, the application of **system mapping** (see Box 9) can be useful.⁴³ The visual aspect of system mapping puts complex concepts and relationships into simpler pictorial representations.⁴⁴ An example of system mapping is presented in Figure 10, showing a value chain as a system and presenting its sub-systems and their boundaries (drawn from the 2019 Corporate-level Evaluation on IFAD's Engagement in Pro-poor Value Chain Development). This mapping was instrumental to present the intricacy of working on value chain development, particularly when trying to generate equitable outcomes for smaller producers.

Box 9

Systems mapping

For evaluating higher-level programs characterized by significant causal complexity, system mapping is particularly valuable. System mapping provides a structured approach to identifying and presenting the systemic nature of programs embedded in their contexts. The primary purpose of system mapping is to describe the different components of a system and how these are connected. Mapping makes complex systems more comprehensible, therefore more approachable. There are a number of ways to approach mapping the system to represent system elements and connections. This include: **actor maps** (to show which individuals and/or organizations are key players in the space and how they are connected); **causal-loop** diagrams (to clarify the positive and negative feedback loops that lead to system behavior or functioning); **Issue maps** (to lay out the political, social, or economic issues affecting a given geography or constituency); **mind maps** (to highlight various trends in the external environment that influence the issue at hand); social network analysis (to understand the social structures and networks operating within the system) and many others.

In evaluation, system mapping is particularly relevant for understanding, for example, the institutional, social, economic, and cultural aspects of the context in which a program operates and how they influence how the program works. This supports a better understanding of the nature and impact of complex programs.

To ease the process, new online tools are also becoming available, several of which combine multiple mapping methods. Some of the more popular tools include: [Insight Maker](#) and [Kumu](#)

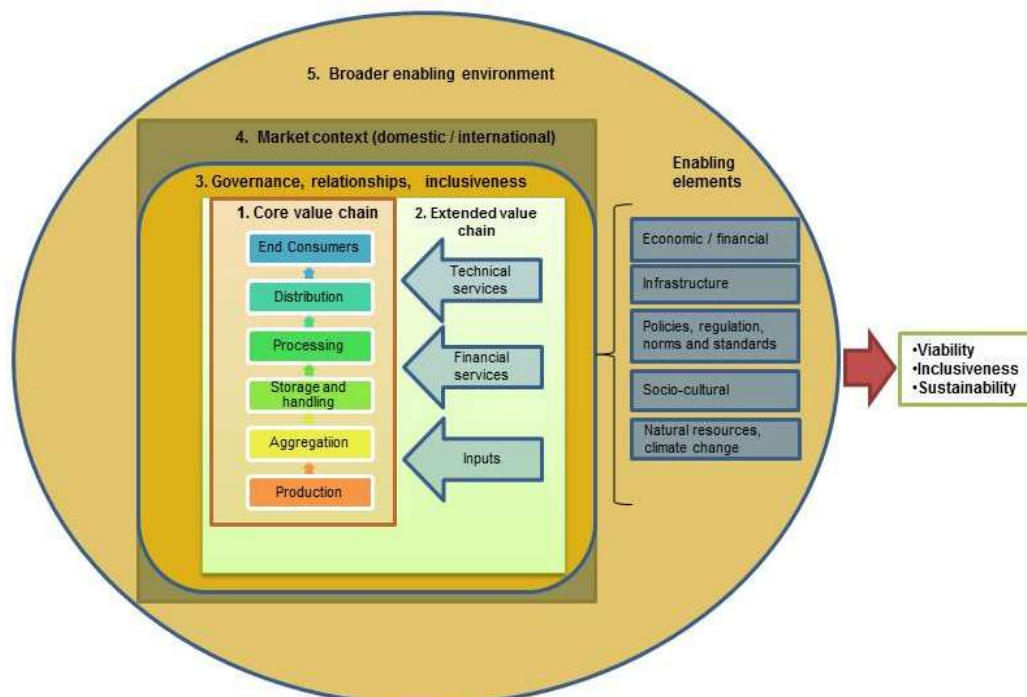
Examples:

- Barbrook-Johnson, P. and Penn, A. (2021) 'Participatory systems mapping for complex energy policy evaluation', *Evaluation*, 27(1), pp. 57–79.
- Participatory systems mapping: exploring and negotiating complexity in evaluation with BEIS and Defra. A presentation on how system mapping can be a highlight participatory process.

⁴³ Jos Vaessen, Sebastian Lemire, and Barbara Befani, *Evaluation of International Development Interventions: An Overview of Approaches and Methods*; Washington, DC . Independent Evaluation Group World Bank, 2020.

⁴⁴ For a practical example, see: Barbrook-Johnson, P. and Penn, A. "Participatory systems mapping for complex energy policy evaluation", *Evaluation*, (2021): 27(1), pp. 57–79.

Figure 10
Mapping of a value chain system and its sub-systems (Corporate-level Evaluation Value Chain, 2019)



Source: Corporate-level Evaluation adaptation from FAO (2014), with inputs from GIZ (2018) and USAID (2014).

IFAD's attribution analysis approach

66. As highlighted in Chapter I and Chapter II, IFAD conducts evaluations assessing the contribution to overall change with [the dual purpose of learning and accountability](#). Nonetheless, all evaluations also aim at detecting to what extent a particular outcome or impact can be "attributed" to a given intervention. However, this is not a methodologically and operationally easy assessment. First, IFAD-supported activities involve many partners therefore disentangling the exact impact attributable to IFAD's intervention might not be straightforward. Second, beneficiaries are exposed to external factors that influence results, particularly donor countries' policies, beneficiary countries' domestic policies, other development programmes, socio-economic fluctuations, structural changes and climatic phenomena. Third, baselines that may provide strong support for dealing with attribution issues are often not available or not of the required quality. Therefore, making a robust assessment attributing the results achieved on the ground to a particular intervention is challenging and expensive.
67. Given these challenges, theory-based design with a combination of methods for data collection, analysis and triangulation is the overarching approach to evaluation in IFAD, as discussed in more detail in the following section. At the same time, IFAD has a unique approach to attribution analysis through both independent and self-evaluations.
68. IOE conducts impact evaluations on selected completed operations with the main aim to: (i) establish more solid evidence basis for future strategic evaluations; (ii) test innovative methodologies for assessing the results of IFAD interventions more rigorously; and (iii) contribute to the ongoing internal and external debate on the measurement of the impact of development interventions.
69. IFAD management's systematic attribution analysis is conducted by RIA, which designs impact assessments to specifically detect the attributable impact of IFAD's interventions. This approach entails conducting impact assessments at the project level on a representative sample of 15 per cent of IFAD's operation portfolio. The

data is then aggregated and projected to the whole portfolio at corporate level, addressing the challenges to establishing attribution with a rigorous quantitative methodology. This process complements the traditional independent and self-evaluation approaches with a systematic approach to attribution analysis. In addition, Management reports the results measured through the Core Outcome Indicators (COI) in the project log frame,⁴⁵ which are also obtained through a rigorous survey methodology to establish attribution. This evaluation process makes IFAD unique among the IFIs in doing that.

70. The quantitative methodology is based on economic theory and empirical literature on impact evaluations using non-experimental ex-post methods.⁴⁶ It starts with the selection of a representative sample of projects closing in one IFAD replenishment period. Once the representativeness of the sample is confirmed, the methodology includes: i) creating the ToC of the project; (ii) creating a robust sample frame that includes beneficiaries and a carefully selected control group (i.e. that represent the counterfactual); (iii) collecting quantitative and qualitative data from both groups (around 2,000–3,000 households); (iv) conducting analysis using non-experimental methodologies for each of the selected projects to estimate attributable impact on a large set of development indicators (i.e. change in each indicator for beneficiaries compared to a control group); (v) conducting a meta-analysis to estimate overall impact; and (vi) conducting a projection analysis to the whole IFAD portfolio for corporate reporting and learning. In particular, RIA Impact Assessments Report on the COI of IFAD's Strategic Framework 2016-2025 defined in the Results Management Framework of IFAD as Tier II development impact indicators. These are the economic goal of increasing incomes and the three strategic objectives of improving productive capacities, market access and strengthening the environmental sustainability and climate resilience, as well as cross-cutting themes of food and nutrition security and women's empowerment.

C. Evaluation criteria, key questions and ratings

71. In line with good practice in international development evaluation, IFAD uses a set of **evaluation criteria** to assess the performance of policies, strategies, operations and business processes.⁴⁷ The use of evaluation criteria supports consistent, high-quality evaluation across IFAD and facilitates the aggregation of findings to conduct additional thematic analyses (e.g. across regions, topics, type of countries) as well as analysis over time.
72. As shown in Figure 11, in addition to the six internationally established criteria (OECD-DAC),⁴⁸ IFAD adopts further criteria that address its specific mandate.
73. Compared to previous editions, this manual introduces four main changes. The purpose of these changes are: (i) to align with international standards; (ii) to avoid excessive fragmentation and repetition in the structure of reports, encouraging the preparation of more concise documents; (iii) for Management to follow reporting practices agreed with Member States in the context of IFAD11 and IFAD12.
74. First, the Manual introduces the "coherence" criterion (now part of the international reference criteria) to be used mainly for country and corporate/thematic

⁴⁵ The RIMS and associated log frame guidance was replaced in 2017 with Core Indicators (CIs). In 2017, with the migration of all logframes from paper/PDF to ORMS, all RIMS indicators were converted into CIs into all log frames for ongoing projects. Core Outcome Indicators (COIs) are a sub-set of CIs, and will become mandatory for all projects designed starting from 2022. Results for COIs are captured through surveys carried out at three times over the course of project implementation: at project baseline, mid-term and completion stages. A specific, IFAD-tailored methodology has been developed for these surveys and is found in the [Core Outcome Indicators measurement guidelines](#).

⁴⁶ For a reference literature of the method used in this approach see: Rubin (1974), Rosenbaum and Rubin (1983), Heckman et al. (1997), Hahn (1998), Heckman et al. (1998), Dehejia and Wahaba (1999), Wooldridge (2007), Hirano et al. (2003), Caliendo and Kopeinig (2008), Imbens and Wooldridge (2009), Wooldridge (2010), Austin (2011).

⁴⁷ Apart from evaluation product of RIA, all other evaluations use the mentioned evaluation criteria and system of ratings.

⁴⁸ Relevance, coherence, effectiveness, efficiency, impact and sustainability.

See: <https://www.oecd.org/dac/evaluation/daccriteriaforevaluatingdevelopmentassistance.htm>

- evaluations.⁴⁹ The “non-lending activities” (Knowledge Management, Partnership Building and Policy Dialogue) are assessed as subdomains of coherence in evaluations at the country, corporate and thematic levels.
75. Second, the Manual now merges the two criteria of “environment and natural resources management” and “climate change and adaptation” in one single criteria. This is in line with the Intergovernmental Panel on Climate Change (2018) definition of climate resilience: “In human systems, the process of adjustment to actual or expected climate and its effects, in order to moderate harm or exploit beneficial opportunities. In natural systems, the process of adjustment to actual climate and its effects”. In IFAD’s criteria, adaptation is therefore not seen in isolation. The key goals of adaptation strategies are to build the resilience of people and agricultural systems to climate change and to sustain and enhance the livelihoods of poor people. These strategies consequently need to be rooted in an understanding of how the poor and vulnerable sustain their livelihoods, the role of natural resources in livelihood activities and the scope for adaptation actions that reduce vulnerabilities and increase the resilience of poor people. For the purpose of self-evaluation, and only for projects closing until the IFAD12 period, the two dimensions related to “environment and natural resources management” and “climate change and adaptation” will continue to receive separate ratings at self-evaluation stage, to allow reporting on RMF11 and RMF12 targets agreed with Member States. The two ratings will be aggregated for the purpose of IOE reporting.
76. Third, the Manual presents a new arrangement of the criteria. “Innovation” is assessed under effectiveness, whereas “Scaling-up” and “Environment and natural resources management and climate change adaptation” are assessed under sustainability.⁵⁰ IOE will provide individual ratings for these IFAD-specific criteria. While management will not rate “innovation”, it will provide an individual rating to “Scaling-up”, “Environment and natural resources management” and “climate change adaptation” only for projects closing until IFAD12, to allow reporting on RMF11 and RMF12 targets agreed with Member States.
77. Although Management will not rate these criteria by the six point rating scale, it will monitor and measure them in COSOPs and at project completion and implementation. Management is re-defining the approaches to these issues in IFAD 12: the IFAD12 matrix of commitments includes an action plan on sustainability, and a scaling up strategy to be completed in 2022. The approach to natural resource management and climate change is captured in the revised SECAP guidelines (2021, being rolled out in 2022) and in the relevant COIs and mandatory CIs in the annex III (page 54). Turning to innovation, the IFAD 12 monitorable action n.26 aims at developing an operating model and guidelines for innovation to be led by the Change Delivery and Innovation Unit; this is also to be finalised in 2022. These 2022 actions are to refine further the measurement approaches. Eventually, and in consultation with IOE, these objective measurements can form the basis for management ‘rating’ these criteria.
78. Fourth, Management and IOE will follow different courses of action with regard to impact. IFAD Management assesses impact using a rigorous approach in five domains, in line with those included in Tier II development indicators of its Results Management Framework (RMF):
- (i) Economic mobility;

⁴⁹ <https://www.oecd.org/dac/evaluation/daccriteriaforevaluatingdevelopmentassistance.htm>

⁵⁰ Knowledge management, partnership building and policy dialogue (assessed under coherence), innovation (assessed under effectiveness), scaling-up and “environment and natural resources management and climate change adaptation” (assessed under sustainability), will continue to be rated individually. Part 2 of the Manual provides guidance for assigning overall ratings at the project or country programme level.

- (ii) SO1 Productive capacities (agricultural/non-agricultural production and productivity);
- (iii) SO2 Access to market (access and integration into markets);
- (iv) SO3 Resilience (ability of households to cope with climate and non-climate shocks);
- (v) Improved nutrition.
79. As mentioned, Management measures impact through rigorous Impact Assessments undertaken on 15 per cent of the portfolio, which rely on a counterfactual-based analysis and guarantees attribution of results, thus representing a robust measure of the impact of IFAD-supported projects. In addition, Management reports the results measured through the COIs in the project logframe, which are also obtained through a rigorous survey methodology to establish attribution. Starting in IFAD12 all projects will carry out COIs surveys at baseline, midterm and completion stages (with treatment and comparison groups) since this is the best way to capture the results of a project's interventions over the course of its implementation. The Core Outcome Indicators Measurement Guidelines (2021) provide the COI surveys methodology. Table A in Annex III presents the full list of COIs, mapped to the respective impact domain. Through the RMF, Management reports on the impact of its projects at aggregated level. Given that it presents quantitative evidence on attributable impact of projects based on rigorous methodologies, starting from IFAD12, Management does not rate impact.
80. IOE will continue to rate impact, according to the international practice. IOE will draw evidence from impact studies undertaken by projects, by IFAD Management, or other organizations and validate the findings independently, based on available information, fieldwork, and its own expertise. In selected cases, IOE may conduct its own impact surveys. Similarly, IOE will take into account the data available through the COI surveys, validate them as above and will collect additional data and information as required by the specific operations, development context or independent evaluation questions.

Figure 11

IFAD's evaluation criteria

Project-level Evaluations	Country-level Evaluations
<p>International criteria</p> <p>Relevance</p> <p>Effectiveness</p> <ul style="list-style-type: none"> • Innovation* <p>Efficiency</p> <p>Impact *</p> <ul style="list-style-type: none"> • Changes in: incomes and assets; social/human capital; households food security and nutrition; institution and policies <p>Sustainability</p> <ul style="list-style-type: none"> • Scaling-up* • Natural resource management and climate change adaptation* <p>IFAD-specific</p> <p>Gender equality and women empowerment</p> <p>Performance of partners</p> <ul style="list-style-type: none"> • IFAD • Government 	<p><i>Criteria used for project-level evaluations and, in addition, the international criterion of Coherence</i></p> <ul style="list-style-type: none"> • Non-lending activities <ul style="list-style-type: none"> ✓ Knowledge Management ✓ Partnership Development ✓ Policy Engagement

* These criteria will continue to be rated by IOE, not by Management. With regard to scaling-up and natural resource management and climate change, Management will only rate these criteria for projects closing until the IFAD12 period (2022-2024), in order to comply with RMF11 and RMF12 reporting requirements.

81. Table 1 presents the definitions of the criteria and the related overarching evaluation questions. The use of core questions help ensure consistency and

comparability. It allow aggregation of ratings across IFAD evaluations and helps focus data collection and analysis.

Table 1
Definition of the evaluation criteria used by IFAD

<i>Evaluation criteria (project and country levels)</i>	<i>Overarching questions</i>
<p><u>Relevance</u> The extent to which: (i) the objectives of the intervention/ strategy are consistent with beneficiaries' requirements, country needs, institutional priorities and partner and donor policies; (ii) the design of the interventions / strategy*, the targeting strategies adopted are consistent with the objectives; and (iii) the intervention / strategy has been (re-) adapted to address changes in the context.</p> <p>*Evaluations will analyse the strategy pursued whether explicit (written) or implicit.</p>	<p>Was the intervention/country strategy and programme relevant and aligned to:</p> <p>(a) the country's development needs and challenges as well as national policies and strategies; (b) IFAD's relevant strategies and priorities; (c) the needs of the beneficiaries and tailored to very poor or marginalized people or special categories.</p> <p>Was the design quality in line with available knowledge, recognized standards (if available)?</p> <p>Was the design realistic in terms of meeting the context and implementation capacity?</p> <p>Was the design re-adapted to changes in the context (if applicable)?</p>
<p><u>Coherence</u> (mainly for country level and strategic evaluations)</p> <p>This comprises the notions of external and internal coherence. The external coherence is the consistency of the strategy with other actors' interventions in the same context. Internal coherence looks at the internal logic of the strategy, including the complementarity of lending and non-lending objectives within the country programme.</p> <p><u>Non-lending activities</u> are specific domains to assess coherence.</p> <p><u>Knowledge management</u> The extent to which the IFAD-funded country programme is capturing, creating, distilling, sharing and using knowledge.</p> <p><u>Partnership building</u> The extent to which IFAD is building timely, effective and sustainable partnerships with government institutions, international organizations, private sector, organizations representing marginalized groups and other development partners to cooperate, avoid duplication of efforts and leverage the scaling up of recognized good practices and innovations in support of small-holder agriculture and rural development.</p> <p><u>Policy engagement</u> The extent to which IFAD and its country-level stakeholders engage, and the progress made, to support dialogue on policy priorities or the design, implementation and assessment of formal institutions, policies and programmes that shape the economic opportunities for large numbers of rural people to move out of poverty</p>	<p>What is the overall coherence of the country programme? In particular:</p> <ul style="list-style-type: none"> To what extent were there synergies and interlinkages between different elements of the country strategy/programme (i.e. projects, non-lending activities)? How did IFAD position itself and its work in partnership with other development partners? Did IFAD contribute to policy discussion drawing from its programme experience? To what extent lessons and knowledge have been gathered, documented and disseminated?
<p><u>Effectiveness</u> The extent to which the intervention/country strategy achieved, or is expected to achieve, its objectives and its results at the time of the evaluation, including any differential results across groups</p> <p>A specific sub-domain of effectiveness relates to <u>Innovation</u>, the extent to which interventions brought a solution (practice, approach/method, process, product, or rule) that is novel, with respect to the specific context, time frame and stakeholders (intended users of the solution), with the purpose of improving performance and/or addressing challenge(s) in relation to rural poverty reduction.⁵¹</p>	<p>Were the objectives of the intervention/country strategy and programme achieved or likely to be achieved at the time of the evaluation?</p> <p>Did the intervention / strategy achieve other objectives or did it have any unexpected consequence?</p> <p>To what extent did the programme or project support / promote innovations, aligned with stakeholders' needs or challenges they faced? Were the innovations inclusive and accessible to a diversity of farmers (in terms of gender, youths, and diversity of socio-economic groups)?</p>

⁵¹ Conditions that qualify an innovation: newness to the context, to the intended users and the intended purpose of improving performance. Furthermore, the 2020 Corporate-level Evaluation on IFAD's support to Innovation defined transformational innovations as "those that are able to lift poor farmers above a threshold, where they cannot easily fall back after a shock". Those innovations tackle simultaneously multiple challenges faced by smallholder farmers. In IFAD

Efficiency

The extent to which the intervention or strategy delivers, or is likely to deliver, results in an economic and timely way “Economic” is the conversion of inputs (e.g., funds, expertise, natural resources, time) into outputs, outcomes and impacts, in the most cost-effective way possible, as compared to feasible alternatives in the context. “Timely” delivery is within the intended timeframe, or a timeframe reasonably adjusted to the demands of the evolving context. This may include assessing operational efficiency (how well the intervention was managed).

Having considered the nature of intervention and implementation context, key questions include (but are not limited to):

What is the relation between benefits and costs (e.g., net present value, internal rate of return)? How does it compare with similar interventions (if the comparison is plausible)?⁵²

Are unit costs of specific interventions in line with recognised practices and congruent with the results achieved?

Are programme management cost ratios justifiable in terms of intervention objectives, results achieved, considering contextual aspects and unforeseeable events?

Is the timeframe of the intervention development and implementation justifiable, taking into account the results achieved, the specific context and unforeseeable events?

Impact

The extent to which an intervention/country strategy has generated or is expected to generate significant positive or negative, intended or unintended, higher-level effects.

The criterion includes the following domains:

- changes in incomes, assets and productive capacities
- changes in social / human capital
- changes in household food security and nutrition
- changes in institution and policies

The analysis of impact will seek to determine whether changes have been transformational, generating changes that can lead societies onto fundamentally different development pathways (e.g., due to the size or distributional effects of changes to poor and marginalized groups)

Has the intervention/country strategy and programme had the anticipated impact on the target group and institutions and policies? Why?

What are the observed changes in incomes, assets of the target group, household food security and nutrition, social/human capital and institutions and policies over the project/COSOP period? What explains those changes? What are the challenges?

From an equity perspective, have very poor / marginalized groups, special categories, benefited in a sizable manner?

Sustainability

The extent to which the net benefits of the intervention or strategy continue and are scaled-up (or are likely to continue and be scaled-up) by government authorities, donor organizations, the private sector and others agencies.

Note: This entails an examination of the financial, economic, social, environmental, and institutional capacities of the systems needed to sustain net benefits over time. It involves analyses of resilience, risks and potential trade-offs.

Specific domain of sustainability:

Environment and natural resources management and climate change adaptation. The extent to which the development interventions/strategy contribute to enhancing the environmental sustainability and resilience to climate change in small-scale agriculture.

Scaling-up* takes place when: (i) bi- and multi laterals partners, private sector, communities) adopt and diffuse the solution tested by IFAD; (ii) other stakeholders invested resources to bring the solution at scale; and (iii) the government applies a policy framework to generalize the solution tested by IFAD (from practice to policy).

*Note that scaling up does not only relate to innovations

To what extent did the intervention/country strategy and programme contribute to long-term institutional, environmental and social sustainability?

What is the level of engagement, participation and ownership of the government, local communities, grass-roots organizations and the rural poor? In particular, did the government ensure budget allocations to cover operation and maintenance?

Did the programme include an exit strategy?⁵³

- For Environment and natural resources management and climate change adaptation, to what extent is the intervention / strategy:

-Improving farming practices? Minimizing the damage and introducing offsets to counter the damage caused by those farming practices?

-Supporting agricultural productivity that is sustainable and integrated into ecosystems?

-Channelling climate and environmental finance through the intervention/country programme to smallholder farmers, helping them to reduce poverty, enhance biodiversity, increase yields and lower greenhouse gas emissions?

-Building climate resilience by managing competing land-use systems while reducing poverty, enhancing biodiversity, increasing yields and lowering greenhouse gas emissions?

- For scaling up:

operation contexts, this happens by packaging / bundling together several small innovations. They are most of the time holistic solutions or approaches applied of implemented by IFAD supported operations.

⁵² References to Management documents that relate to this criterion include: (i) the IFAD action plan of efficiency; (ii) the IFAD Internal Guidelines on Economic and Financial Analysis of rural investment Projects, 2015; and (iii) IFAD’s project implementation Guidelines, Annex VII - Value for Money in Supervision.

⁵³ Useful references to Management’s documents related to this criterion include the IFAD action plan on sustainability and the IFAD Project Design Guidelines, 2020 (notably annex V).

-To what extent were results scaled up or likely to be scaled up in the future? Is there an indication of commitment of the government and key stakeholders in scaling-up interventions and approaches, for example, in terms of provision of funds for selected activities, human resources availability, continuity of pro-poor policies and participatory development approaches, and institutional support?

Gender equality and women's empowerment.

The extent to which IFAD interventions have contributed to better gender equality and women's empowerment. For example, in terms of women's access to and ownership of assets, resources and services; participation in decision making; workload balance and impact on women's incomes, nutrition and livelihoods; and in promoting sustainable, inclusive and far-reaching changes in social norms, attitudes, behaviours and beliefs underpinning gender inequality.

Evaluations will assess to what extent interventions and strategies have been gender transformational, relative to the context, by: (i) addressing root causes of gender inequality and discrimination; (ii) acting upon gender roles, norms and power relations; (iii) promoting broader processes of social change (beyond the immediate intervention).

Evaluators will consider differential impacts by gender and the way they interact with other forms of discrimination (such as age, race, ethnicity, social status and disability), also known as gender intersectionality.⁵⁴

What were the project's achievements in terms of promoting gender equality and women's empowerment, including intersectionality issues?

In particular, were there changes in: (i) women's access to resources, income sources, assets (including land) and services; (ii) women's influence in decision-making within the household and community; (iii) workload distribution (including domestic chores); (iv) women's health, skills, nutrition?

Were there notable changes in social norms, attitudes, behaviours and beliefs and policies / laws relating to gender equality?

Was attention given to programme implementation resources and disaggregated monitoring with respect to gender equality and women's empowerment goals?

Performance of partners (assessed separately for IFAD and the Government)

The extent to which IFAD and the Government (including central and local authorities and executing agencies) supported design, implementation and the achievement of results, conducive policy environment, and impact and the sustainability of the intervention/country programme

The adequacy of the Borrower's assumption of ownership and responsibility during all project phases, including government and implementing agency, in ensuring quality preparation and implementation, compliance with covenants and agreements, supporting a conducive policy environment and establishing the basis for sustainability, and fostering participation by the project's stakeholders.

Performance of IFAD

How effectively did IFAD support the overall quality of design, including aspects related to project approach, compliance, and operational aspects?⁵⁵

How proactively did IFAD identify and address threats to the achievement of project development objectives?⁵⁶

How effectively did IFAD support the executing agency on the aspects of project management, financial management, and setting-up project level M&E systems?

How did IFAD position itself and its work in partnership with other development partners?

Performance of the Government

How tangible was the Government's commitment to achieving development objectives and ownership of the strategy / project?

Did the Government adequately involve and consult beneficiaries/stakeholders at design and during implementation?

How did the Government position itself and its work in partnership with other development partners?

How well did the Project Management Unit/Project Coordination Unit manage start up procedures, implementation arrangements, appointment of key staff, and resource allocation/funding?

How timely did the project management unit (PMU) identify and resolve implementation issues? Was project management responsive to changes in the environment or the recommendations made during supervision missions or by the Project Steering Committee?

⁵⁴ Evaluation Cooperation Group, "Gender. Main messages and findings from the ECG Gender practitioners' workshops) (Washington, D.C., 2017), <https://www.ecgnet.org/document/main-messages-and-findings-ieq-gender-practitioners-workshop>

⁵⁵ Useful reference to Management's comments that relate to the criterion are the IFAD project design guidelines (2020).

⁵⁶ Sources for self-evaluations include Project Supervisions and Project Status Reports.

How useful were the various project management tools (AWPB) and the Management Information System (MIS) developed during implementation? Were these tools properly used by project management?

How did the PMU fulfil fiduciary responsibilities? How useful was the procurement plan and how was it used during implementation?

How adequate were monitoring and evaluation arrangements made by the PMU, including the M&E plan, and the utilization of evaluation M&E data in decision-making and resource allocation?

Ratings

82. Evaluation criteria are scored according to a rating system introduced by IFAD in 2002. In 2005, IFAD moved from a four-point to a six-point rating system in line with the practice adopted in many other IFIs and United Nations organizations,⁵⁷ allowing for a more nuanced assessment of project results. In addition to reporting on performance based on the six-point rating scale, in 2007 IFAD introduced the broad categories of "satisfactory" and "unsatisfactory" for reporting on performance across the various evaluation criteria (see **table 2**).⁵⁸
83. As a general rule, evaluators assign ratings, supporting their arguments with evidence and justifying the ratings with solid analysis. Evaluators are often faced with what is called the 'hindsight issue'. This refers to the challenge of assessing (and rating) past projects with current perspectives, notably when applying evaluation criteria. Box 7 presents a way forward for dealing with the hindsight issue.

Box 7

Evaluators are often faced with the challenge of assessing (and rating) past projects with current perspectives, notably when applying evaluation criteria. This is particularly challenging when the political or operational context, the operation targets and/or strategies have changed during the course of implementation. Holding managers accountable for failing to achieve today's standards before they were known may be unfair. For example, in cases where context or policies have changed late in the life of a project without the opportunities for course correction, *prima facie* it may seem to be anachronistic to assess managerial performance with today's metrics.

At the same time, it is a common fact that the context within which projects and country programmes are implemented can change. Change in complex systems is characterized by uncertainty, volatility, and adaptation. In order to perform, a project or strategy needs to be able to adapt, not just to stick rigidly to its original formulation, when the context changes or is no more conducive. Therefore, in several cases, it may be possible for programme managers to adapt to the context. In such cases, it is legitimate for an evaluation to assess the extent to which a programme has been resilient to the change in context or has been adapted to respond to changes.

All the above requires a balancing act in an evaluation. On the one hand, it is fair to expect some capacity to adapt to changes in context. On the other hand, the evaluators need to acknowledge when standards have changed significantly in recent times and cannot be applied retroactively.

Dealing with changes in contexts and standards over time (the hindsight issue)

⁵⁷ Evaluation Cooperation Group, *ECG Good Practice Standards for Evaluation of MDB Supported Public Sector Operations*. For each rated criterion, Multilateral Development Banks use an even number (mostly four, exceptionally six, for greater differentiation) of rating scale points. For the sake of validity, credibility, transparency and comparability they apply a clearly defined rating for each scale point that reflects a pre-defined set of ranked value terms"

⁵⁸ The Annual Review of Development Effectiveness produced by the Independent Evaluation Group of the World Bank uses a similar system of categorization.

84. Ratings for individual criteria are given using integers (i.e. no decimal points).⁵⁹ Consistent with most other evaluation offices and to keep the system simple, no weights are assigned to ratings when determining a final rating for overall project achievement. This Manual provides the following general guidance (table 2) to support evaluators to assign ratings on each evaluation criteria. With the aim of further reducing the space for subjectivity in ratings, as well as the disconnect between self- and independent evaluation, part 2 of this manual will provide further guidance on ratings.

Table 2
Rating system at IFAD

Score	Assessment	Indicative description	Category
6	Highly Satisfactory	The activity, project, programme, policy achieved the large majority of the main targets, objectives, expectations, results (or impacts). Due to its high quality, it could be considered as an 'outstanding practice' case. A rating of '6' may also signal that results (e.g. effectiveness, impacts) are 'transformational'	Satisfactory
5	Satisfactory	The activity, project, programme, policy achieved the large majority of the main targets, objectives, expectations, results (or impacts). Regarding its quality, the intervention / policy in question could be considered as a good practice example that inspire other programmes / policies.	
4	Moderately Satisfactory	The activity achieved the relative majority of the targets, objectives, expectations, results or impacts. At the same time, there were some notable gaps in achievement. The quality of what was achieved was good, although not a special case of good practice.	
3	Moderately unsatisfactory	Under the concerned criterion, the activity did not achieve the relative majority of its targets and objectives, results (or impact). There were areas of clear under-achievement. In qualitative terms, achievements were below standards and expectations	Unsatisfactory
2	Unsatisfactory	Under the concerned criterion, the activity achieved only a minority of its targets, objectives, expectations, results or impacts. Quality of achievement was low and well below standards	
1	Highly unsatisfactory	Under the concerned criterion, the activity (project, programme, non-lending) achieved almost none of its targets, objectives, expectations, results or impacts. Quality was very poor and there may have been cases of worsening of the situation.	

Source: Elaboration by this Manual (2021).

D. Reviewing evaluability and data availability

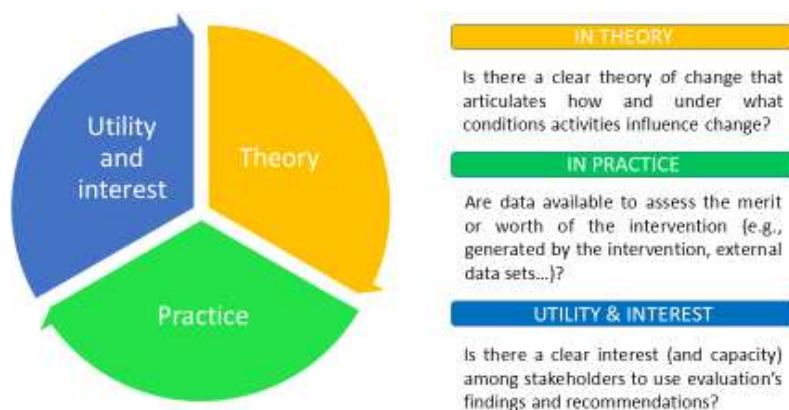
85. In the evaluation literature, the notion of **evaluability assessment** is related to an analysis to be conducted before deciding whether a specific evaluation should be undertaken and when it should be undertaken.⁶⁰ A graphic example is displayed in Figure 12.

Figure 12

What do evaluability assessments examine?

⁵⁹ Evaluations may establish composite ratings (e.g. arithmetic averages of other ratings) which would be rational numbers, with decimals. As an example, a rating for the overall project achievements could be established as the arithmetic average of all the individual project ratings.

⁶⁰ OECD-DAC, Glossary of key terms in evaluation and results-based management (Paris: OECD-DAC, 2010), p. 21.

**Examples:**

- Evaluability Assessment Checklist, developed by Davies, R (2013) *Planning Evaluability Assessments: A Synthesis of the Literature with Recommendations*. Report of a Study Commissioned by the Department for International Development.
- ILO Evaluation Office (2020) Evaluability assessment checklist.

Source: Adapted from Betterevaluation.org.

86. However, the reality in many development organizations is that the decision to carry out a specific evaluation is often driven by governing bodies and corporate-level commitments. Thus, the assessment of what can be evaluated often happens after an evaluation has been approved. As already discussed at the beginning of this chapter, the examination of the scope, approach and evaluability are in part overlapping and iterative processes. Thus, it is important to recognize that the review of evaluability may have a feedback loop on the scope of an evaluation.
87. For many evaluations, the review of evaluability and data availability typically focuses on:
- whether it is timely to assess the results of a policy, strategy or project, or whether the evaluation should be confined to the on-going process and the likely pathway towards achieving certain results;
 - whether secondary data are available to complete the analysis according to specific evaluation criteria and what are the main information gaps to be filled in and through what methods; and
 - to what extent it will be possible to collect and analyse certain data, depending on time, budget and other circumstances (e.g., security, credibility, social acceptance).
88. Thus, the assessment of evaluability and data availability is an important consideration when taking decisions on data collection and analysis. Relevant inputs include the review of background documentation and databases (e.g. World Bank, United Nations system, think tanks, literature), official documentation from IFAD, government and other agencies, surveys, preliminary interviews with the main stakeholders (both in person and virtual). In some cases, a brief reconnaissance mission to a country or project site may be required to complete the review.

E. Approaches and methods for data collection and analysis

89. Data collection tools vary according to the type of evaluation scope, approach, outcome of the review of evaluability, availability of secondary data and other contextual factors.
90. **Mixed methods for analytical rigour and depth.** In line with international good practice,⁶¹ IFAD encourages triangulation of methods, data collection and data analysis. All evaluations must be based on evidence and need to explicitly consider limitations related to the analysis conducted. Evaluators will always strive to

⁶¹ Evaluation Cooperation Group, *ECG Big Book on Good Practice Standards* (Washington, D.C., 2012).

identify and use the best-suited methods for the specific purposes and context of the evaluation and consider how other methods may compensate for any limitations of the selected methods. In particular, strategic and aggregate-level evaluations – e.g. multi-level, multi-site evaluations in country, thematic and cluster evaluations – are by definition multi-method. However, the idea of informed evaluation design, or the strategic mixing of methods applies to all evaluations.

91. Using a mix of designs and methods and triangulating information from different approaches is recommended to assess different facets of complex outcomes or impacts as well as to capture the cultural and contextual complexities that affect the achievement of the desired goals.^{62 63} This yields greater validity than a single method.⁶⁴
92. Keeping **cultural responsiveness** at the forefront of evaluation efforts is a key component to broadening participation and incorporating culture and context into an evaluation. This requires evaluators to be sensitive and responsive to the cultural context in which the programme and/or policy is operationalized at all stages of the evaluation process. This is particularly important in instances when the participants' culture is known to have a major influence on outcomes. Ensuring a systematic and coherent application of culturally responsive evaluation begins with integrating cultural dimensions into the evaluation framework (Box 11).⁶⁵ Since culturally responsive evaluation is an emerging approach, UNEG has identified key questions that could be posed as a minimum, and aspirational questions that can be considered when designing and implementing an evaluation.⁶⁶

Box 11

A quote on culturally responsive evaluation

"Culturally responsive evaluations are based on the notion that evaluation cannot be separated from the sociocultural contexts within which programmes are implemented. Culture shapes the behaviours and worldviews of its members and is therefore central to our understanding of individuals' motivations, attitudes and responses to an intervention[...] To be 'culture-blind' in evaluation runs the risk of perpetuating inequalities, in the same way that 'gender-blind' evaluation or policy does".

Source: UNEG (2018) Development of culturally responsive criteria for evaluations.

93. For evaluations to adopt participatory⁶⁷ and culturally responsive approaches,⁶⁸ understanding the context and engaging the stakeholders are key steps for

⁶² C. Bolinson, D.M. Mertens, *Transformative evaluation and impact investing: A fruitful marriage*. In R.P. Herman & E. de Morais Sarmiento (Eds.). *Global Handbook of Impact Investing* (Wiley, 2021).

⁶³ Hur Hassnain, Lauren Kelly and Simona Somma, eds., *Evaluation in Contexts of Fragility, Conflict and Violence: Guidance from Global Evaluation Practitioners* (Exeter, UK: IDEAS, 2021).

⁶⁴ For example, UN Women evaluated its contribution to gender responsive budgeting in Europe and Central Asia by combining participatory theory of change approach with outcomes harvesting – an approach for capturing unintended positive, and negative outcome-level results of interventions within complex contexts. To strengthen the depth and utilization of the analysis, the evaluation also attempted to calculate a Social Return on Investment for Gender Responsive Budgeting. In addition, this evaluation also enriched the analysis by applying, for example, stakeholder mapping using a sequenced application of Critical Systems Heuristics, Human Rights Role Analysis, and Power Analysis; mapping the history of gender responsive budgeting in the region, including a rich picture (systems) view of what forces were at play in determining the decisions that were made; and an institutional pathways analysis assessing the system dynamics that have influenced the history of the GRB in the region and frame the options for future change. See example: UNWOMEN (2017) Evaluation of UNWOMEN contribution to gender responsive budgeting in the Europe and Central Asia Region.

⁶⁵ M. Bryan, A. Lewis, A. Culturally Responsive Evaluation as a Form of Critical Qualitative Inquiry. *Oxford Research Encyclopedia of Education* (2019).

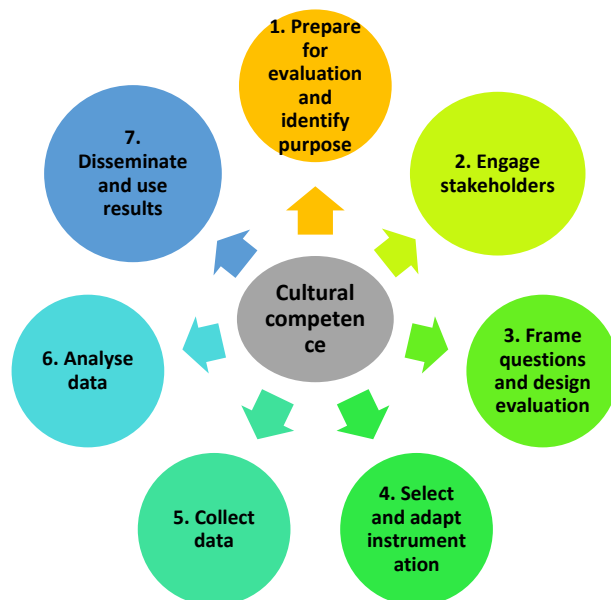
⁶⁶ For a list of culturally responsive evaluation questions and approaches see: UNEG Development of culturally responsive criteria for evaluations (2018), <http://www.unevaluation.org/document/detail/2123>.

⁶⁷ For examples and documents that discuss good practices for participatory methods in evaluation see <https://www.participatorymethods.org/>

⁶⁸ For more information and examples of culturally responsive evaluation see: J.A Chouinard and F. Cram *Culturally Responsive Approaches to Evaluation* (USA: Sage Publications, 2020); D. Mertens, *Mixed Methods Design in Evaluation* (USA: Sage Publications, 2018); UNEG, *Compendium of Evaluation Methods Reviewed Volume 1* (New York, 2020); Frey, B. *Culturally Responsive Evaluation* (2018); and the Sage Encyclopedia of Educational Research, Measurement, and Evaluation, vol. 4. A recent application is also in B. Chilisa and D. Mertens, "Indigenous Made in Africa Evaluation Frameworks: Addressing Epistemic Violence and Contributing to Social Transformation," *The American Journal of Evaluation*, vol. 42(2) 241-253 (2021).

preparing the entire evaluation process, including data collection, analysis and disseminating results (Figure 13). The most appropriate way of approaching primary stakeholders will mostly depend on the local dynamics, socio-economic settings and customs. Evaluators need to understand the context in which the evaluation activities take place and adapt accordingly, placing culture and the community at the center of the evaluation.

Figure 13
Culturally Responsive Evaluation Framework



Source: Adapted from Frierson et al. (2002) and Hood et al. (2015), quoted in Chouinard and Cram (2018).

94. **An integrated gender focus is also fundamental to the collection of relevant information.** In addition to strengthening validity through triangulation of different data collection methods, mixed methods are particularly important for gender and equity responsive evaluation to: (i) study processes of empowerment and behavioral change that are difficult to capture with a single data collection method; (ii) strengthen generalizability of in-depth qualitative analysis (i.e. to ensure the sample of respondents is representative of the total sample population).⁶⁹
95. Adopting participatory methods and data collection tools for evaluation and systematically developing evaluation frameworks that include the voice of marginalized people are a key component of IFAD's evaluations. Not only does this help accurately collect the voice of underrepresented groups but also increases the validity and reliability of the evaluation. Participation can occur at any stage of the evaluation process: in its design, data collection, in analysis, in reporting; they are not exclusive to specific evaluation methods or restricted to quantitative or qualitative data collection and analysis.⁷⁰ The needs and decisions about the type and extent of participation are usually different for an evaluation that focuses on local-level impacts from those of an evaluation that examines national-level change. IFAD evaluations need to pay attention to promoting participation of key stakeholders in the evaluation process, but at the same time ensure that the principles of impartiality, credibility and transparency of the evaluation's analysis and final judgements are upheld.
96. Figure 14 presents commonly used data collection methods. The list is not exhaustive, and a specific evaluation product might need a particular data collection

⁶⁹ Evaluation Cooperation Group, *Integrating gender into project-level evaluation*. ECG reference document. (Washington, D.C., 2017).

⁷⁰ I. Guijt, *Participatory Approaches, Methodological Briefs: Impact Evaluation 5* (Florence: UNICEF Office of Research, 2014).

method that is not included in the list. Most evaluations use a combination of methods, as needed.

97. In the context of strategic and aggregate-level evaluations, designs may encompass different case study levels, with cross-case (comparative) analysis across countries (or interventions). Case studies are often used in IFAD evaluation although they present a number of methodological challenges, particularly in terms of internal and external validity. In terms of internal validity, the concern is how to ensure quality, reliability and robustness of methods and design. The concern with external validity is generalizations, i.e. the extent to which it is possible to generalize, and in which circumstances. Lastly, there are issues relating to aggregation and synthesizing for learning purposes. There is an ongoing international debate on this and different ways to analyse and synthesize findings. One such approach is Qualitative Comparative Analysis (see Box 12).⁷¹

⁷¹ V. Pattyn, A. Molenvend, B. Befani, "Qualitative Comparative Analysis as an Evaluation Tool: Lessons From an Application in Development Cooperation," *American Journal of Evaluation*, vol. 40(1) 55-74 (2019).

Box 12

Qualitative comparative analysis⁷²(QCA)

Primarily designed to answer the questions: (i) Under what circumstances did the program generate or not generate the desired outcome? and/or (ii) What works best, why and under what circumstances? In essence, QCA is a case-based method that enables evaluators to systematically compare cases which are responsible for the success or failure of an intervention by identifying key factors in each case. What differentiates this approach from most other cross-case comparative methods is that it provides a specific set of algorithms to analyse data sets (usually in the form of a table) by using Boolean algebra logic operators to document varying configurations of conditions associated with observed outcomes. In this sense, QCA can also be considered a data analysis technique.

QCA is usually designed for use with an intermediate number of cases – typically between 10 and 50 cases. It is not appropriate in all circumstances as it requires a strong theory of change and clearly defined cases and cannot measure the net effects of an intervention.



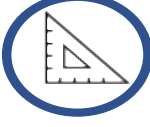




Example:

- Impact Evaluation of the Global Environment Facility Support to Protected Areas and Protected Area Systems (September 2016). In this evaluation, the evaluators used a theory-based design combining multiple methods, including multi-level analysis (global and portfolio) and qualitative comparative analysis. Available [here](#).

Source: Valérie Pattyn, Astrid Molenveld, Barbara Befani, *Qualitative Comparative Analysis as an Evaluation Tool: Lessons from an Application in Development Cooperation* (2017).

⁷² Valérie Pattyn, Astrid Molenveld, Barbara Befani *Qualitative Comparative Analysis as an Evaluation Tool: Lessons from an Application in Development Cooperation*, (2017)

Figure 14
Frequently used data collection methods

		WHEN TO USE
	DOCUMENT REVIEW	To identify available data by reviewing formal policy documents, M&E reports, programme records, political, socio-economic agricultural profiles of the country or specific locale.
	DIRECT OBSERVATION	To learn how the programme naturally occurs by observing sites, practices, living conditions, physical constructions using a well-design observation record form (notes, photos or video)
	PHYSICAL MEASUREMENT	To measure physical changes based on agreed indicators and measurement procedures. Examples include birth weight, nutrition levels, rain levels, and soil fertility.
	INTERVIEWS	To understand individual experiences in more detail. Can be unstructured, semi-structured or structure questions.
	SURVEY	To collect information from a defined group. They are standardize instruments and are usually comprised of well-defined, close ended questions. Can be administered in person, mail, telephone
	FOCUS GROUP DISCUSSION	To discover issues that are of most concern for a community or group
	CASE STUDY	To examine in-depth a limited number of cases. Useful for documenting contextual conditions and producing insights about whether the program might make a difference in other settings
	MEMORY RECALL	To reconstruct beneficiaries and other stakeholders, situation before the project.
	SYSTEMATIC REVIEW	To gather all available empirical data by using clearly defined, systematic methods to obtain answers to specific question.
	EVIDENCE GAP MAPS	To identify key “gaps” where few or no evidence from impact evaluations and systematic reviews is available

Note: This illustration provides a selection of available methods. For a rapid review of methods, see, for example, Vaessen, et al. (2020), IEG World Bank, op cit.

F. The Role of Information and Communication Technologies (ICTs) for data collection and analysis

98. Evaluators across all regions of the world face recurring challenges in the field. Lack of reliable monitoring and evaluation data, limited time and resources, and operating in contexts that are often fragile and affected by conflict and violence are some of the more common obstacles. The outbreak of COVID-19 pandemic and related travel restrictions have adversely affected the design, implementation and evaluation of international development interventions, and the ability of the evaluation function to capture the consequences of the economic crisis facing the rural poor and marginalized people.⁷³ New technologies for data collection and analysis (and new types of data) are slowly but steadily making their entry into the practice of international development and its evaluation. This is an area of increasing interest for IFAD. In 2017, IOE organized an international conference on ICT for evaluation and published a book on the subject.⁷⁴ In 2019, IFAD prepared a Strategy for Information and Communication Technology for Development.⁷⁵
99. The increasing emphasis on complexity, real-time feedback and adaptive management approaches (see [Chapter II](#)), coupled with the COVID-19 pandemic, have accelerated remote data collection to minimize the risk of spreading the virus and underscored the urgency of getting accurate data quickly. ICTs are offering new methods and tools for gathering, analyzing and disseminating data, and are changing the way evaluations are conducted, potentially opening the door to more rigorous evaluation.
100. A variety of tools is now available for evaluators that enable more data to be collected, often remotely, and to be processed faster. A comprehensive description of the vast array of emerging technologies for data collection is beyond the scope of the manual, but links to other sources of information are provided where relevant.⁷⁶ Figure 15 presents a summary of the most prominent tools and methods for data collection and analysis. Different tools offer specific strengths and weaknesses. Typical opportunities offered by ICT-inspired innovations for evaluation are for data collection and data analysis. There are also opportunities for data and information display and communication activities.





⁷³ https://www.ifad.org/documents/38714182/42217951/LearningNote_Covid19_forweb2.pdf/98f22bb0-6c22-16c3-c54b-4f09b4f0fdcd?t=1610977391000

⁷⁴ O. Garcia, P. Kotturi, eds., *Information and Communication Technologies for Development Evaluation* (Routledge, 2019).




⁷⁵ <https://webapps.ifad.org/members/eb/128/docs/EB-2019-128-R-5.pdf>

⁷⁶ For a detailed discussion of the role of ICTs and big data in evaluation practice see, for example: P. York, M. Bamberger, *Measuring results and impact in the age of big data: the nexus of evaluation, analytics, and digital technology* (New York: The Rockefeller Foundation, 2020); and Hassnain, H., Kelly, L., Somma, S., eds. "Evaluation in Contexts of Fragility, Conflict and Violence Guidance from Global Evaluation Practitioners" (IDEAS, 2021).

Figure 15
ICT tools: Advantages and Disadvantages⁷⁷

ICTs for data collection			
	Description	Advantages	Disadvantages
 MOBILE DATA COLLECTION	The targeted gathering of structured information using mobile phones, tablets or PDAs using a special software application	It can improve the timeliness and accuracy of the data collection. Platforms allow customization of survey to include photographs, voice recordings, GPS coordinates	Technology alone will not improve the survey design or instrument. Potential bias in favour of well-educated or well-off citizens
 REMOTE SENSING	Observing and analyzing a distant target using electromagnetic spectrum of satellites, aircrafts or other airborne devices	Possible to collect data on inaccessible areas. Observed objects or people are not disturbed.	Privacy concerns Potentially high costs for obtaining images or for primary data collection using remote sensors. Socio-economic indicators hard to capture.
 GEOGRAPHIC INFORMATION SYSTEMS	Computer-based tool for integrating and analyzing geographic or spatial data.	Combination of different types of geographical data sets. It allows viewing, interpreting and visualizing data into numbers of ways –revealing relationships, trends and patterns. GIS can also be used to digitally represent and interpret oral and life histories and can accommodate qualitative information.	GIS setup is complex. In addition to the cost of the equipment, there is training cost. Frequent updating of datasets or data models may lead to errors in results.
 CLOUD COMPUTING	Delivery of computing services – servers, storage, databases, networking, software, analytics and more – over the Internet (“the cloud”), thus enabling shared access to resources.	Access to data storage and analytical tools in a shared manner enables organizations to operate effortlessly across geographical areas Real-time integration of data collection, analysis and reporting.	Security concerns over access to data. Requires robust and high-speed internet connection.
ICTs for Data Analysis			
Description	Advantage	Disadvantage	

⁷⁷ Multiple Source, see for instance: UNDP, *Discussion Paper: Innovations in Monitoring & Evaluating Results* (New York, 2013). INTRAC, *ICT in Monitoring and Evaluation* (Oxford, 2017). O. Garcia, P. Kotturi, eds., *Information and Communication Technologies for Development Evaluation* (Routledge, 2019).

	Representation of data graphically and interactively	Graphical and interactive presentation of data increases accessibility of complex data sets and, in turn, the use of the data. It can identify trends and patterns of complex and large data sets.	Identifying and putting together data visualization can be time-consuming, or costly if Outsourced.
DATA VISUALIZATION			
	A set of methods for getting computers to recognize words and images, and creating prediction models	Review large volumes of data and identify patterns, trends or specific information.	Needs large data sets to train on. Initial algorithm training is time consuming.
MACHINE LEARNING			
	The use of advanced analytic techniques against very large, diverse big data sets that include structured, semi-structured and unstructured data, from different sources, and in different sizes.	Access to a range of descriptive, exploratory and predictive analytics tools, which makes it possible to develop models and for evaluating complex programs and predict future trends.	"Data exhaust" is not representative of the wider population, much less of the marginalized.
BIG DATA ANALYTICS			

Source: Elaboration by the authors (2021).

101. ICTs offer an unprecedented number of options for evaluators to access, gather and analyse data more efficiently. ICTs enable evaluators to go further in exploring the ToC (see [Section 3.3](#)), and to do so with greater rigour. For that reason, they are critical to strengthen evidence-based policymaking that relies on evaluation findings. Evaluators need to keep abreast of what ICT tools are available in order to best decide when and how to incorporate them into their work.⁷⁸
102. However, **ICTs are not a panacea**, but a means to an end. Technology is only as good as the evaluators who use it; and evaluations of development programmes will still need to be grounded in robust theory. ICTs can also risk increasing biases, where assumptions included into the computing models can lead to a false sense of objectivity regarding the results. Evaluators must be aware of inherent biases that may be built into the data collection and coding processes and/or the software used to analyse the data.
103. Furthermore, the introduction of ICTs cannot be seen in a stand-alone manner but as a part of an organization-wide process. ICTs must be mainstreamed into IFAD's operations. This may include mainstreaming technology into planning, monitoring and evaluation, and self-assessment processes.

⁷⁸ O. Garcia, P. Kotturi, eds. *Information and Communication Technologies for Development Evaluation* (Routledge, 2019).

104. Issues concerning data privacy, ethics and inclusiveness relating to the use of ICTs for evaluation will need to be taken into consideration. In IFAD, the following principles will drive the selection of ICT tools for evaluation:⁷⁹
- **People-centric:** keeps the interest of target groups at the centre of the use of ICTs. Evaluations will not use technology solely for the sake of innovation.
 - **Inclusive:** serves the task of including vulnerable and marginalized populations in the evaluation process. The issue of power discrepancy between those who produce the data and those who use it is vital in this context.
 - **Mixed-methods:** combine traditional, participatory face-to-face data gathering with technology-enabled data collection methods and larger-scale data analytics. This addresses concerns about inclusiveness, makes sense of what big data patterns are showing, and what might be missing from big data sources, and ensures that important contextual clues are not missed.
 - **Privacy and ethics:** protecting privacy and following ethical guidelines in how information is collected and shared to make sure beneficiaries are not put at risk.

G. Evaluation conclusions and recommendations

105. Each evaluation should clearly present **conclusions** in the form of key messages that are informed by the main findings but are not a repetition or a simple summary of the findings. Conclusions bring findings in the report to a “higher-level.” They add value to the findings by providing an answer to the over-arching questions of the evaluation. They also provide explanation to the findings, highlighting the main underlying factors.
106. Conclusions help bridge the findings and the recommendations. However, conclusions should be kept separate from the recommendations, both in content and language (e.g. conclusions should not state what ought to be done in order to improve a certain situation).
107. Conclusions are more forceful when they concentrate on a limited number of judgment-statements (indicatively three to six) that take into consideration the overall findings of the report and point to the main learning items from the evaluation: what worked; what did not work; and what were the key factors. This helps transition to the recommendations.
108. **Recommendations** are proposals of actions made to entities in charge of a programme, a strategy, and/or policies to bring about improvements in performance and results. The quality of **recommendations** is a crucial factor of the evaluation to optimally stimulate learning, accountability and organizational effectiveness. UNEG Improved Quality of Evaluation Recommendation Checklist (2018) defines evaluation recommendations as “*proposals aimed at enhancing the effectiveness, efficiency, impact, relevance, sustainability, coherence, added value or coverage of the operation, portfolio, strategy or policy under evaluation. Recommendations are intended to inform decision making, including programmed design and resources allocations.*”⁸⁰ To this end, developing recommendations involves weighing effective alternatives, policy, and funding priorities within a broader context. It requires in-depth contextual knowledge, particularly about the organizational context within which policy and programme decisions will be made and the political, social and economic context in which investments operate.
109. Care must be taken to ensure that recommendations are: (i) appropriate for achieving the objectives of the interventions; (ii) are few in numbers (typically from

⁷⁹ These principles draw from several sources. In particular, L. Raftree, L., *Technology, Biases and Ethics: Exploring the Soft Sides of Information and Communication Technologies for Evaluation* (ICT4Eval, 2019), in O. Garcia, P. Kotturi, (eds), op.cit.

⁸⁰ UNEG “Improved Quality of Evaluation Recommendation Checklist”. *Working paper* (2018). This checklist includes useful background information, and it also deals with the follow-up of evaluation recommendations, using the UNEG Good Practice Guidelines for Follow-up.

three to six); (iii) positioned strategically; and (iv) once implemented, will add value to the organization. Recent guidance on preparing recommendation is available from the UNEG's Improved Quality of Evaluation Recommendations Checklist (2018) (Box 13) and from ECG Practice Note Formulation of Evaluation Recommendations (2018).⁸¹

Box 13

UNEG quality standards for recommendations

- The report describes the process followed in developing the recommendations, including consultation with stakeholders.
- Recommendations are firmly based on evidence and conclusions.
- Recommendations are relevant to the objectives and purposes of the evaluation.
- Recommendations clearly identify the target group for each recommendation.
- Recommendations are clearly stated, with priorities for action made clear.
- Recommendations are actionable and reflect an understanding of the commissioning organization and potential constraints to follow-up.

Source: UNEG Quality Assurance Tool for the *Formulation, Structure and Content* of Recommendations (2018).

110. As highlighted earlier in this manual, the full utility of an evaluation hinges on participation, dissemination, learning and follow-up. Therefore, recommendations need to be presented in a form that allows different decision-makers to clearly identify their responsibility. This should also facilitate tracking of follow-up actions by IFAD's Management in the annual President's Report on the Implementation Status of Evaluation Recommendations and Management Actions (PRISMA). IFAD Management is also planning to establish an on-line system to track recommendations.
111. Part 1 (Chapters I-III) of this manual has provided the overall context for evaluation in IFAD, addressing IFAD's evaluation foundational elements, including its mission, evaluation objectives, architecture, frameworks, principles, and criteria that guide all evaluations within IFAD. Part 2 provides practical and detailed guidance on different evaluation products covering both independent and self-evaluations, as well as the linkages between them.

⁸¹ Evaluation Cooperation Group, ECG Practice Note Formulation of Evaluation Recommendations (Washington, D.C., 2018).

IFAD Internal Evaluation Architecture

Self-evaluation system

1. At the core of self-evaluation is the Development Effectiveness Framework (DEF). Introduced in 2016, the DEF was developed to ensure that evidence is collected from projects and is systematically used; and to create the necessary structure to facilitate the collection and use of evidence in decisions regarding the design and implementation of projects. Self-evaluation projects are designed with the purpose of achieving the expected results of the DEF, namely, to strengthen accountability, enhance learning, and ultimately ensuring that the decision-making process is fully based on reliable evidence.⁸²
2. To ensure relevance to the IFAD12 business model, which is centered on transformational country programmes and supported by institutional change and a revised financial framework, an updated version of the DEF will be adopted, starting 2022. The updated DEF provides the framework for improving IFAD's self-evaluation structure around three key pillars: (i) enhancing monitoring, evaluation, adaptation and learning; (ii) enhancing capacity, mainstreaming, sustainability, efficiency and scaling-up; and (iii) working at the country level to maximize impact beyond projects. In line with the updated DEF, and IFAD12 commitments, IFAD is developing dedicated action plans for areas where IOE and Management have consistently seen that project/programme performance has been weak: sustainability, efficiency, scaling up, monitoring, evaluation, adaptation and learning. Updates will also be undertaken in areas such as working in conditions of conflict and fragility, and IFAD's offer on country-level policy engagement.
3. Self-evaluation products are developed at three main levels: country level, project level, and corporate level.
4. **At the country / COSOP level**, self-evaluation starts at design, when the results framework for the country strategy is reviewed by IFAD's Quality Assurance Group, Operational Policy and Results Division (OPR) and other members of the Operational Strategy and Policy Guidance Committee utilizing the Development Effectiveness Matrix for COSOPs. Light touch reviews are conducted every year, and halfway through implementation. COSOP Results Reviews are undertaken to assess progress towards results, lessons learned, risk factors encountered and changes in country demand and priorities. At completion, COSOPs undergo a completion review, i.e. a self-evaluation of their strategic objectives and IFAD's performance in achieving them. Lessons learned from IFAD engagement feed the preparation of new COSOPs.
5. **At the project level**, self-evaluation is fully integrated into the operation life cycle. At design, the Development Effectiveness Matrix is used to review and enhance evaluability. To this purpose, the operation's ToC and logical framework are reviewed, including impact, outcome and output indicators, together with their baseline and target values.
6. During implementation, project teams prepare the annual Supervision Report, by describing progress achieved and identifying the main challenges encountered during execution. They also update progress data against logframe indicators and targets, and rate project performance according to a set of pre-defined criteria. Following an adaptive management approach, such information is used at project level to identify corrective actions and adjust the annual work plan and budget, including through the creation of a Project Improvement Plan if needed. At portfolio-management level, the information from the logical frameworks and project supervision reports are used to ensure that adequate expertise and budget are allocated where performance needs to be followed upon or corrected. At mid-term, project teams conduct a full stock take of progress achieved and report it in the Mid-Term Review; according to the evidence collected, the logframe is updated

⁸² <https://webapps.ifad.org/members/ec/115/docs/EC-2021-115-W-P-6.pdf>

- as needed and relevant decisions regarding the future of the operation are made, including possible restructuring.
7. At the end of the operation execution period, the relevant regional division prepares a Project Completion Report (PCR). Through the PCR, project teams (under previous practice, government actors/the PMU) rate the operation development effectiveness according to the standard criteria, and additional ones in line with IFAD strategies, including: rural poverty impact; environment and natural resource management; climate change adaptation; and gender equality. PCRs also include a section on lessons learned, to benefit the design of new operations and improve implementation by building on experience.
 8. In addition to these common self-evaluation practices that are applied to all projects, RIA conducts rigorous impact assessments of a representative sample of approximately 15 per cent of the projects closing in each replenishment period. Impact assessments use non-experimental methods to estimate the attributable impact of individual projects on IFAD strategic goal and objectives. Qualitative methods are also used to provide additional information on the context to complement the analysis. IFAD's impact assessment agenda makes up an important component of self-evaluation at the project and corporate levels. Furthermore, individual project impact assessments enrich the PCR in evaluating the rural impact.
 9. Under the updated DEF, and in line with IFAD's graduation policy,⁸³ COSOP guidelines are being updated and improved to foster better tools for adaptation and learning, and enhance country-level sustainability, scalability, partnership and policy influence. In addition, project supervision guidelines are being revised with the integration of tools for better data collection and monitoring, with special attention to the use of geographic information systems/satellite data to enhance M&E systems. Moreover, the existing guidelines to collect core outcome indicators⁸⁴ are increasingly being integrated into project design and will have a special emphasis in the revised supervision guidelines. PCR guidelines are also being updated to reflect a more objective scoring scale as well as to benefit more from RIA impact assessments (both in terms of data and estimated attributable impacts) in cases where the project falls in the impact assessment sample. The responsibility for preparing the PCR is being shifted from governments to IFAD, which is expected to improve PCR quality, candour, timeliness, and transparency.
 10. **At the corporate level**, the Report on IFAD's Development Effectiveness presents the annual operational and organizational performance, by reporting on a set of 79 Results Management Framework indicators agreed upon with Member States. Additionally, IFAD undertakes thematic or cluster reviews on areas that are of specific interest to the organization. It uses data from ongoing projects on a continuous basis through stocktaking to place emphasis on areas/countries/regions in which performance requires attention. This emphasis on real-time data contributes to IFAD's culture of results beyond specific self-evaluation products.
 11. In addition, the attributable impacts estimated by RIA for a sample of projects are aggregated in a meta-analysis at the end of each replenishment period. This then feeds into a projection exercise to calculate attributable impact of IFAD's overall portfolio during that period. The results contribute to corporate reporting and learning to improve future design and targeting for better impact.
 12. Information on most self-evaluation products is monitored and captured through online systems; each one dedicated to a specific purpose and stage of the project life cycle, from concept to ex-post evaluation (see Table 1). Data from self-evaluation products are also presented in dedicated dashboards both for internal and external audiences.

⁸³ <https://webapps.ifad.org/members/eb/133/docs/EB-2021-133-R-5.pdf>

⁸⁴ A set of indicators that measure the change expected, as a result of participation of beneficiaries in the project, collected through a rigorous methodology that demonstrates attribution of results.

Table 1
IFAD Data systems

Operation Document Center (ODC)	Internal corporate system for sharing and managing operation documents produced in all IFAD interventions. It is designed to manage project and programme documents from design through completion, as well as documents across global/corporate, regional and country portfolios.
Grants and Investment Projects System (GRIPS)	Internal corporate system for keeping a record of projects financed through investment or grant programmes, together with their financial information.
<u>Operational Results Management System (ORMS)</u>	Internal system for the management and tracking of quantitative and qualitative project information related to: logframe indicators – baselines, targets and progress data; performance during implementation; development effectiveness at completion; action tracker; and lessons learned.
<u>Quality Assurance Archiving System (QUASAR)</u>	Internal platform for managing the quality assurance review of all project designs and soon, of all grants, concept notes and COSOPs.
Commitment Tracker	Internal tool used to track monitorable actions and outputs to fulfill commitments taken under IFAD's replenishment. First formulated for IFAD11, it is being replicated for commitments agreed upon for IFAD12 (2022-2024).
Operations Dashboard	Internal dashboard that provides up-to-date information on the performance of IFAD-supported country strategies and projects/programmes. Displays data on design, implementation, and performance rating (with most external data published on IFAD website).
RMF Dashboard	Section on IFAD website where all donors, stakeholders and interested parties can view up-to-date progress towards IFAD's targets for the 79 indicators that Member State's selected to track during the IFAD11 period. It is being adapted for IFAD12.

Independent evaluation

13. Independent evaluations are conducted by the Independent Office of Evaluation (IOE), which is structurally, functionally, and behaviourally independent from IFAD Management. IOE ensures that the whole evaluation function at IFAD follows internationally recognized good standards and practices. Independent evaluations help reveal what has been achieved, what does or does not work, and guide the development of successful policies and strategies to support rural transformation. The target audience of independent evaluations are IFAD's Management and Governing Bodies, Member Countries and the international development community at large.
14. IOE conducts a range of independent evaluations at different levels, including project, country program, sectoral, thematic and corporate.
15. Project-level evaluations include the independent validations of PCRs, project performance evaluations, impact evaluations and project cluster evaluations (the latter examine a set of projects in different countries that have a common topic of concentration, for example, rural finance). These products inform higher-plane evaluations, as well as the design of new and ongoing operations.
16. Country-level and regional evaluations include CSPEs and sub-regional evaluations. CSPEs are usually conducted before IFAD and the concerned government prepare a new results-based COSOP and, as such, their findings and recommendations feed into the design of new COSOPs. Sub-regional evaluations assess intra-regional issues or common development challenges within the region, aligned with IFAD's decentralized business model.
17. Project- and country-level evaluations are the building blocks for evaluation syntheses. Evidence from past evaluations is synthesized and analysed to present evaluative knowledge on topics of strategic relevance, and to inform future directions and corporate-level evaluations. Corporate-level evaluations generate lessons and recommendation to enhance IFAD's future policies and strategies. More

detailed information and specific processes for each evaluation product are presented in Part 2 of this manual.

18. Finally, the IOE's Annual Report on Results and Impact of IFAD Operations presents a synthesis of the performance of IFAD-supported operations, and highlights systemic and cross-cutting issues, lessons and challenges that emerge from all independent evaluations. It also identifies recommendations to enhance IFAD's development effectiveness. From 2022, IOE will produce the Annual Report on Independent Evaluation, a revamped version of its annual report.

Evaluation that are responsive to social justice and gender equality

1. **Social justice and nobody left behind.** The commitment of Agenda 2030 to “leaving no one behind”⁸⁵ needs to be reflected in evaluations. Evaluators should assess whether: (i) programmes have undertaken an analysis of the inequalities between different groups; (ii) the underlying drivers of such inequalities; (iii) whether programme designs address such inequalities; and (iv) whether results frameworks of interventions have indicators to measure the progress.⁸⁶ In line with a human rights-based approach, evaluations should assess the extent to which the initiative has facilitated the capacity of rights-holders to claim their rights and duty-bearers to fulfil their obligations.⁸⁷
2. Typical example of discrimination and power imbalance include: (i) economic factors (income, wealth); (ii) ethnicity (also indigenous status in some countries); (iii) social categories (including castes in some countries); (iii) gender; (iv) belonging to political groups/factions; (v) belonging to religious groups; and (vi) health and disability. This list is not comprehensive, and the evaluators may identify other sources of discrimination and imbalance.
3. Evaluators need to assess how inclusive the intervention has been for different beneficiary groups and how key principles such as equity, non-discrimination and accountability have been incorporated from design to results.⁸⁸ There is a need for balancing aggregation versus specificity, with stronger focus on the categories left behind, and on the “last mile” project delivery, rather than on average coverage and results. During data collection and data analysis, it is essential to consider to what extent needs and priorities are being addressed. An analysis of differential results across groups and the extent to which the intervention contributes to or exacerbates equity gaps is a critical element for evaluations.⁸⁹
4. **Gender equality as a specific area of attention for IFAD.** In line with the Fund’s mandate, policies, strategies and work experience, evaluations at IFAD aim to be gender-responsive. It does so by providing a systematic and impartial assessment that delivers credible and reliable evidence-based information about the extent to which an intervention has resulted in progress towards intended and/or unintended results regarding gender equality and the empowerment of women. IFAD evaluations need to assess the degree to which gender and power relationships (including structural and other causes that give rise to inequities, discrimination and unfair power relations), change as a result of an intervention using a process that is inclusive, participatory and respectful of all stakeholders (rights-holders and duty-bearers).⁹⁰

⁸⁵ The 2030 Agenda emphasizes the importance of empowering people who are vulnerable, including children, youth, persons with disabilities, people living with HIV/AIDS, older persons, indigenous peoples, refugees and internally displaced persons and migrants.

⁸⁶ See United Nations Sustainable Development Group, “Leaving No One Behind: A UNSDG Operational Guide for UN Country Teams” (interim draft) (2019). <https://unsdg.un.org/resources/leaving-no-one-behind-unsdg-operational-guide-un-country-teams>

⁸⁷ See United Nations Evaluation Group, Norms and Standards for Evaluation. New York: UNEG. Human rights and gender equality are considered a norm (Norm 8 on human rights and gender equality) and a standard (Standard 4.7, “The evaluation design should include considerations of the extent to which the United Nations system’s commitment to the human rights-based approach and gender mainstreaming strategy was incorporated in the design of the evaluation subject.”) (2016) <http://www.unevaluation.org/document/detail/1914>

⁸⁸ United Nations Evaluation Group, “Guidance on Evaluating Institutional Gender Mainstreaming” (New York: UNEG, 2018), <http://www.unevaluation.org/document/detail/2133>

⁸⁹ Evaluators should ensure consistency and accuracy of terminology used in relation to gender issues in official documentation and publications, following the 2017 IFAD Glossary on gender issues: IFAD, *Glossary on gender issues*. (2017), <https://www.ifad.org/en/web/knowledge/-/publication/glossary-on-gender-issu-1>

⁹⁰ Evaluation approaches must integrate gender equality concerns and are all subject to assessment against the United Nations System-wide Action Plan on Gender Equality and the Empowerment of Women (UN-SWAP) [Evaluation Performance Indicator](#). Likewise, all evaluations are subject to assessment against the [United Nations Disability Inclusion Strategy](#) (UNDIS), and the inclusion of persons with disability should be considered in all phases of the evaluation process and in every type of evaluation (*include hyperlink to IOE guidance note or UNEG guidance note - forthcoming).

5. At IFAD, performance in relation to gender equality and women’s empowerment assessed against a scale, moving progressively from “gender blind” (i.e. there were no attempts to address gender concerns and/or the result had a negative outcome; aggravated, or reinforced existing gender inequalities and norms), through “gender mainstreaming” (i.e. gender equality and women’s empowerment have been mainstreamed, such that all three strategic objectives of the IFAD gender policy have been addressed),^{91 92} all the way to, at the top of the scale, gender transformative (i.e. gender power dynamics have been transformed by addressing social norms, practices, attitudes, beliefs and value systems that represent structural barriers to women’s and girls’ inclusion and empowerment). What “gender transformative change” means depends on the context (Box 1). Different benchmarks are needed for different contexts and good contextual analysis is a general prerequisite.

Box 1.

Definition of gender transformative approaches⁹³

Gender transformative approaches are defined as those that aim to overcome the root causes of inequality and discrimination through promoting sustainable, inclusive and far-reaching social change. They actively seek to transform gender power dynamics by addressing social norms, practices, attitudes, beliefs and value systems that represent structural barriers to women’s and girls’ inclusion and empowerment. They seek to ensure equal access for women to productive assets and services, employment and market opportunities, and supportive national policies and laws. Transformation and entry points towards it are context-specific and take into account that women are not a homogeneous group.

Source: Evaluation Synthesis on gender equality (2017).

6. In assessing women’s empowerment, evaluators may refer to the “domains of empowerment” outlined in the IFPRI/FAO/IFAD Women’s Empowerment in Agriculture Index guidelines, namely: decisions about agricultural production; access to and decision-making power about productive resources; control of use of income; leadership in the community, and time allocation.^{94 95} Evaluators may also refer to other analytical frameworks, such as the [gender@work](#) framework to better understand the types of changes that have taken place across the interlinked domains of individual change, formal change, systemic change and informal change.⁹⁶

⁹¹ IFAD, “Policy on Gender Equality and Women’s Empowerment,” (2012),

https://www.ifad.org/documents/38711624/39417906/genderpolicy_e.pdf/dc871a59-05c4-47ac-9868-7c6cfc67f05c?t=1507215182000

⁹² The three strategic objectives of the IFAD Policy on Gender Equality and Women’s Empowerment are: 1) promote economic empowerment to enable rural women and men to have equal opportunity to participate in, and benefit from, profitable economic activities; 2) enable women and men to have equal voice and influence in rural institutions and organizations, and; 3) achieve a more equitable balance in workloads and in the sharing of economic and social benefits between women and men.

⁹³ IFAD, “Evaluation Synthesis: What works for gender equality and women’s empowerment - a review of practices and results), Independent Office of Evaluation (2017), <https://www.ifad.org/en/web/ioe/-/what-works-for-gender-equality-and-women-s-empowerment-a-review-of-practices-and-results>

https://www.ifad.org/documents/38714182/39721405/gender_synthesis_fullreport.pdf/229358bf-f165-4dcd-9c4a-1af4f09ab065?t=1519897485000; IFAD, “Report of the Consultation on the Twelfth Replenishment of IFAD’s Resources: Recovery, Rebuilding, Resilience,” (2020), <https://webapps.ifad.org/members/repl/12/4/docs/IFAD12-4-R-2-Rev-1.pdf?attach=1>

⁹⁴ IFPRI, “Women’s empowerment in agriculture index,” (2012), <https://www.ifpri.org/publication/womens-empowerment-agriculture-index>

⁹⁵ IFAD, “Measuring women’s empowerment in agriculture: a streamlined approach,” (2019), https://www.ifad.org/documents/38714170/39135332/19_Research_n%C2%B019_web.pdf/37a4a6ec-f23b-44da-8dea-8cefab20f295?eloutlink=imf2ifad

⁹⁶ For more guidance see, for example, UN WOMEN, “Good Practices in Gender-Responsive Evaluations,” (2020), <https://www.unwomen.org/-/media/headquarters/attachments/sections/library/publications/2020/good-practices-in-gender-responsive-evaluations-en.pdf?la=en&vs=2431>

Box 2

Example of issues to be explored that relate to gender equality⁹⁷

- Volume and nature of project resources invested in gender equality and women's empowerment activities.
- Specific activities for gender equality and women's empowerment at the design stage.
- During implementation, to what extent did the project: (i) monitor gender-disaggregated outputs to meet gender equality objectives; (ii) adapt implementation to better meet gender equality and women's empowerment objectives; (iii) address and report on gender issues in supervision and implementation support; (iv) systematically analyse, document and disseminate lessons on gender equality and women's empowerment; and (v) engage in policy dialogue to improve gender equality and women's empowerment.
- In addition to others, changes to: (i) women's access to resources, land, assets and services; (ii) women's influence in decision-making; (iii) workload distribution among household members; (iv) women's health, skills, income and nutritional levels; and (v) gender relations within households, groups and communities in the project area.
- Changes in social norms, attitudes, behaviours, beliefs and value systems that represent structural barriers to women's and girls' inclusion and empowerment; and notice of whether such changes have been reflected in national policies and laws.

⁹⁷ IFAD, "Evaluation Synthesis: What works for gender equality and women's empowerment - a review of practices and results," Independent Office of Evaluation (2017), <https://www.ifad.org/en/web/ioe/-/what-works-for-gender-equality-and-women-s-empowerment-a-review-of-practices-and-results>
https://www.ifad.org/documents/38714182/39721405/gender_synthesis_fullreport.pdf/229358bf-f165-4dcd-9c4a-1af4f09ab065?t=1519897485000;

IFAD, "Report of the Consultation on the Twelfth Replenishment of IFAD's Resources: Recovery, Rebuilding, Resilience,"(2020), <https://webapps.ifad.org/members/repl/12/4/docs/IFAD12-4-R-2-Rev-1.pdf?attach=1>

Core outcome indicators

Table A: List of Core Outcome Indicators (to be collected by projects)⁹⁸

	<i>Core Outcome Indicator</i>	<i>Evaluation criteria for which COI is relevant</i>
Access to natural resources	CI 1.2.1: Households reporting improved access to land, forests, water or water bodies for production purposes	Impact - SO1 Productive capacities (agricultural/non-agricultural production and productivity); - Improved nutrition
Access to agricultural technologies and production services	CI 1.2.3: Households reporting reduced water shortage vis-à-vis production needs CI 1.2.2: Households reporting adoption of new/improved inputs, technologies or practices CI 1.2.4: Households reporting an increase in production	Impact - SO1 Productive capacities (agricultural/non-agricultural production and productivity); - Improved nutrition
Inclusive financial services	CI 1.2.5: Households reporting using rural financial services CI 1.2.6: Partner financial service providers with portfolio-at-risk ≥30 days below 5% CI 1.2.7: Partner financial services providers with operational self-sufficiency above 100%	Impact - SO1 Productive capacities (agricultural/non-agricultural production and productivity); - Economic mobility
Nutrition	CI 1.2.8: Women reporting Minimum Dietary Diversity (MDDW) (RMF 11) CI 1.2.9: Households with improved Knowledge, Attitudes and Practices (KAP)	Impact - SO1 Productive capacities (agricultural/non-agricultural production and productivity); - Improved nutrition
Diversified rural enterprises and employment opportunities	CI 2.2.1: New jobs created (IFAD11) In IFAD12, this indicator will be substituted with IFAD12 RMF indicator: Beneficiaries with new jobs/employment opportunities CI 2.2.2: Supported rural enterprises reporting an increase in profit	Impact - SO2 Access to market (access and integration into markets); CI 2.2.2: Impact - Economic mobility
Rural producers' organizations	CI 2.2.3: Rural producers' organizations engaged in formal partnerships/agreements or contracts with public or private entities CI 2.2.4: Supported rural producers' organizations reporting new or improved services provided by their organization	Impact - SO2 Access to market (access and integration into markets);

⁹⁸ CI Guidelines, October 2021

	CI 2.2.5: Rural producers' organizations reporting an increase in sales	
Rural infrastructure	CI 2.2.6: Households reporting improved physical access to markets, processing and storage facilities	Impact - SO2 Access to market (access and integration into markets);
Environmental sustainability and Climate change	CI 3.2.1: Greenhouse gas emissions (CO2e) avoided and/or sequestered (RMF 11) In IFAD12, this indicator will be substituted with IFAD12 RMF indicator: Tons of Greenhouse gas emissions (tCO2e) avoided and/or sequestered CI 3.2.2: Households reporting adoption of environmentally sustainable and climate- resilient technologies and practices (RMF 11) CI 3.2.3: Households reporting a significant reduction in the time spent for collecting water or fuel	Impact - SO3 Resilience (ability of households to cope with climate and non-climate shocks);
Policy	Policy 3: Existing/new laws, regulations, policies or strategies proposed to policy makers for approval, ratification or amendment	Sustainability - Policy Engagement
Empowerment	IE. 2.1: Individuals demonstrating an improvement in empowerment	Sustainability
Stakeholder Feedback	SF 2.1: Households satisfied with project-supported services SF 2.2: Households reporting they can influence decision-making of local authorities and project-supported service providers	Sustainability

Table B. Mandatory CIs, by project type⁹⁹

Mainstreaming themes and corporate commitments		Related indicators	Use and requirements
Climate Finance	Adaptation	OUTCOME CI 3.2.2: (Number) Percentage of persons/households reporting adoption of environmentally sustainable and climate-resilient technologies and practices CI 3.2.3: (Number) Percentage of persons/households reporting a significant reduction in the time spent for collecting water or fuel	At least one of the following CIs The higher the share of adaptation finance, the more intervention-appropriate indicators may be selected

⁹⁹ CI Guidelines, October 2021

Mitigation	OUTCOME CI 3.2.1: <i>Tons of Greenhouse gas emissions (tCO2e) avoided and/or sequestered.</i>	If Appropriate Mandatory
GENDER Transformative	OUTREACH: disaggregated by sex OUTCOME: <i>CI IE.2.1: Individuals demonstrating an improvement in empowerment</i>	Mandatory Mandatory
Nutrition sensitive	OUTREACH: disaggregated by sex and youth OUTCOME: <i>CI 1.2.8: Percentage of women reporting minimum dietary diversity (MDDW14)</i> <i>CI 1.2.9: Percentage of households with improved nutrition Knowledge Attitudes and Practices (KAP)</i>	Mandatory Mandatory At least 1 Outcome CI mandatory
Youth sensitive Stakeholder Feedback	OUTREACH: disaggregated by sex and youth OUTCOME: <i>CI SF.2.1: (Number) Percentage of households satisfied with project-supported services</i> <i>CI SF.2.2: (Number) Percentage of households reporting they can influence decision-making of local authorities and project-supported service providers</i>	Mandatory Both mandatory in projects logframes approved from December 2020 onwards