
تقرير استعراض الأقران الخارجي الثالث لوظيفة التقييم في الصندوق

الوثيقة: EB 2026/148/R.15

بند جدول الأعمال: 6(ج)

التاريخ: 5 أغسطس/آب 2026

التوزيع: عام

اللغة الأصلية: الإنكليزية

للاستعراض

مراجع مفيدة: سياسة التقييم المعدلة في الصندوق ([EB 2021/132/R.5/Rev.1](#))، استراتيجية التقييم المتعددة السنوات ([EC 2021/114/W.P.8](#))، دليل التقييم ([EB 2022/135/R.29](#))
الإجراء: المجلس التنفيذي مدعو إلى استعراض تقرير استعراض الأقران الخارجي الثالث لوظيفة التقييم في الصندوق.

الأسئلة التقنية:

Veronique Salze-Lozac'h

كبيرة المقيمين السابقة

إدارة التقييم المستقل

المصرف الأوروبي للإنشاء والتعمير

البريد الإلكتروني: vsalzelozach@gmail.com

الصندوق الدولي للتنمية الزراعية - <http://www.ifad.org/ar>

تقرير استعراض الأقران الخارجي الثالث لوظيفة التقييم في الصندوق

الوثيقة: EC 2026/133/W.P.5

بند جدول الأعمال: 6

التاريخ: 15 يونيو/حزيران 2026

التوزيع: عام

اللغة الأصلية: الإنكليزية

للاستعراض

مراجع مفيدة: سياسة التقييم المعدلة في الصندوق ([EB 2021/132/R.5/Rev.1](#))، استراتيجية التقييم المتعددة السنوات ([EC 2021/114/W.P.8](#))، دليل التقييم ([EB 2022/135/R.29](#))
الإجراء: لجنة التقييم مدعوة إلى استعراض تقرير استعراض الأقران الخارجي الثالث لوظيفة التقييم في الصندوق.

الأسئلة التقنية:

Veronique Salze-Lozac'h

كبيرة المقيمين السابقة

إدارة التقييم المستقل

المصرف الأوروبي للإنشاء والتعمير

البريد الإلكتروني: vsalzelozach@gmail.com

**استعراض الأقران الخارجي
لوظيفة التقييم في الصندوق
لعام 2026**

التقرير

مايو/أيار 2026

جدول المحتويات

ii.....	تمهيد
iii.....	النتائج والرسائل الرئيسية لاستعراض الأقران الخارجي لعام 2026
1.....	1- المقدمة والخلفية
1.....	1-1 المقدمة
2.....	2-1 الخلفية – تنفيذ توصيات استعراض الأقران الخارجي لعام 2019
4.....	2- النتائج الرئيسية
4.....	1-2 التقييم العام: وظيفة قوية يتعين أن تكون أكثر استراتيجية لمواكبة الصندوق في المستقبل
5.....	2-2 المصادقية: قوية وموثوقة، مع الحاجة إلى اتباع نهج أكثر استراتيجية
11.....	3-2 الاستقلالية والحوكمة: وظيفة راسخة ومحترمة، مع الحاجة إلى ضمان الحفاظ على الاستقلالية ...
12.....	4-2 الفائدة: قيمة معترف بها، ولكن الاستفادة منها محدودة
15.....	5-2 تكامل التقييم المستقل والتقييم الذاتي: تحسينات مع استمرار بعض التحديات
17.....	6-2 الإدارة والموارد البشرية: الحاجة إلى مزيج أوسع من المهارات
18.....	7-2 بيئة تمكينية للتعليم والتطوير: إمكانات كبيرة لمزيد من التحسين
23.....	3- التوصيات

تمهيد

بدأ المجلس التنفيذي للصندوق استعراض الأقران الخارجي الوارد في هذا التقرير وفقا لسياسة التقييم في الصندوق. وأجرى الاستعراض فريق خارجي يتألف من رؤساء التقييم وفقا للمعايير الدولية.

ويقيم هذا الاستعراض جودة وظيفة التقييم – سواء التقييم الذاتي الذي تقوده الإدارة أو التقييم المستقل الذي يديره مكتب التقييم المستقل. ويبين الاستعراض أن وظيفة التقييم حاليا ذات مصداقية وتحظى بتقدير كبير، مع تنفيذ توصيات استعراض الأقران الخارجي لعام 2019.

ومع ذلك، خلص الاستعراض إلى أن الحفاظ على الأثر وتعزيزه في المستقبل يتطلبان اعتماد نهج أكثر استراتيجية وانتقائية وتوجها نحو التعلم. ويشمل ذلك زيادة التركيز على المواضيع الاستراتيجية، وإصدار مخرجات أكثر دقة من حيث التوقيت، وتعزيز المتابعة، وتحقيق مواءمة أفضل بين التقييم المستقل والتقييم الذاتي، وزيادة التركيز على الفرص والتحديات الناشئة عن تزايد استخدام الذكاء الاصطناعي.

ويقدم هذا التقرير النتائج الرئيسية لاستعراض الأقران الخارجي المتعلقة بالمعايير الأساسية الستة: المصداقية، والاستقلالية، والفائدة، والتكامل بين التقييم الذاتي والتقييم المستقل، والإدارة والموارد البشرية، وتهيئة بيئة تمكينية للتعلم والتطور، تليها توصيات لمواصلة تحسين وظيفة التقييم.

ويعبر أعضاء الفريق عن امتنانهم للمشاركة الواسعة في الاستعراض من قبل جميع أقسام الصندوق والدعم الممتاز المقدم خلال العملية. فاستعراض الأقران يشكل دائما تجربة للتعلم المتبادل، يكتسب من خلالها أعضاء الفريق رؤى ويتعلمون من الممارسات الجيدة المطبقة في المنظمة الخاضعة للاستعراض. ونحن نقدر دعوتنا للاضطلاع بهذا الاستعراض.

سان لونيرو/واشنطن العاصمة/يون

مايو/أيار 2026

Véronique Salze-Lozac'h (الرئيسة)

Geeta Berta

Jörg Faust

Per Øyvind Bastøe (استشاري)

النتائج والرسائل الرئيسية لاستعراض الأقران الخارجي لعام 2026

"تُعدّ وظيفة التقييم قوية، ويجري العمل على تعزيز التقييم الذاتي. وهي لا تعاني من خلل جوهري بل تحتاج إلى التطور لتكون ملائمة للغرض المنشود في المستقبل."

المصادقية

النتيجة

يُنظر إلى وظيفة التقييم في الصندوق - ولا سيما مكتب التقييم المستقل - على أنها ذات مصداقية ومهنية وسلطة معترف بها على نطاق واسع، وتحظى بثقة قوية من الإدارة والمجلس التنفيذي وأصحاب المصلحة الخارجيين. وتُعتبر التقييمات القطرية (تقييمات الاستراتيجيات القطرية والبرامج القطرية) ذات مصداقية وتأثير على نحو خاص.

الرسالة الرئيسية

تمثل المصادقية إحدى نقاط القوة الأساسية في وظيفة التقييم في الصندوق وينبغي الحفاظ عليها، حتى مع إجراء تغييرات على مجالات التركيز والانتقائية والتنفيذ.

وقد تؤدي زيادة استخدام الذكاء الاصطناعي إلى زيادة الكفاءة في عمليات التقييم وإمكانية الوصول إلى أدلة التقييم.

الحوكمة والاستقلالية

النتيجة

تُعدّ البيئة التمكينية العامة لوظيفة التقييم في الصندوق مرضية. وتتواءم سياسة التقييم الحالية في الصندوق مع المعايير الدولية. وينبغي توضيح بعض التفاصيل في ملاحق السياسة (مثل أدوار ومسؤوليات مكتب الفعالية الإنمائية، والعمليات والممارسات المتعلقة بالموارد البشرية لاختيار مدير مكتب التقييم المستقل)، لكن لا حاجة إلى تعديل السياسة نفسها.

وتُعدّ استقلالية مكتب التقييم المستقل راسخة ومفهومة جيدا وليست محل جدال. وينظر أصحاب المصلحة دائما إلى الاستقلالية على أنها عنصر أساسي في قيمة مكتب التقييم المستقل، حيث تمكّنه من طرح مسائل حساسة أو غير مريحة. وقد عملت التحسينات الأخيرة في مجال المشاركة على تعزيز مصداقية التقارير دون تقويض هذه الاستقلالية.

الرسالة الرئيسية

يعمل مكتب التقييم المستقل من موقع يحظى باستقلالية قوية. ويتمثل التحدي في ضمان أن يكفل التطور الحالي للتقييم الذاتي جودة وحسن توقيت التقييمات الذاتية، بالإضافة إلى التكامل الواضح مع التقييم المستقل.

وينبغي أن تساهم الممارسات المتطورة للتفاعل والتكامل بين وظيفتي التقييم بشكل ملموس في تعزيز ملائمة نتائج التقييم وقيمتها التعليمية وأثرها.

الفائدة (الملاءمة والاستخدام)

النتيجة

يُنظر إلى تقييمات مكتب التقييم المستقل في الغالب على أنها مفيدة على مستوى المجلس والإدارة العليا، لا سيما فيما يتعلق بالمساءلة والتفكير الاستراتيجي على المستوى القطري.

أما التقييمات الذاتية فهي أكثر تفاوتًا من حيث الجودة والفائدة، وينبغي أن يركز تحويل هذه الوظيفة على جودة وفائدة عمليات وتقارير التقييم الذاتي.

ومع ذلك، فإن الاستخدام من جانب الموظفين التشغيليين متفاوت، كما أن الاستفادة من التعلم تكون مقيّدة أحيانا بطول وتوقيت وحجم مخرجات التقييم وبقدرة الإدارة وأعضاء المجلس على الاستيعاب.

الرسالة الرئيسية

تظل الفائدة والاستخدام متفاوتين على نطاق المنظمة، فهما قويان على مستوى المجلس والإدارة، لكنهما أكثر محدودية على المستوى التشغيلي، مما يحد من فائدة وتأثير التقييم في المؤسسة.

ويكتسي تحسين التركيز الاستراتيجي، وحسن التوقيت، وتحديد الأولويات وإمكانية الوصول أهمية بالغة لتعزيز الاستفادة وصنع القرار. ويمكن أن يساعد استخدام الذكاء الاصطناعي على معالجة هذه المشكلة الشائعة في معظم المؤسسات المالية الدولية، وذلك من خلال تمكين مختلف أصحاب المصلحة من الوصول السريع وعند الطلب إلى كم كبير من رؤى التقييم.

التكامل بين التقييم الذاتي والتقييم المستقل

النتيجة

أدت إعادة معايرة وإنشاء مكتب الفعالية الإنمائية إلى تقليل التوتر وتحسين التعاون بين التقييم الذاتي ومكتب التقييم المستقل. كما تحسن التعاون بشكل كبير بفضل النهج البناء للتعاون، والعمليات التي تفضل التشاور والتبادل، دون تفويض استقلالية مكتب التقييم المستقل.

ومع ذلك، لا تزال هناك تفاوتات منهجية بين التقييم الذاتي والتقييم المستقل، وتفاوت في جودة التقييم الذاتي وجودة البيانات في الأقاليم، وضعف الترابط بين الوظيفتين، مما يحد من الاتساق والتعلم.

الرسالة الرئيسية

شهد التكامل تحسنا، لكنه لا يزال جزئيا؛ وتبرز الحاجة إلى أدوار أوضح، ومواءمة منهجية أقوى، وتفاعل أكثر منهجية.

مزيج المنتجات والتركيز الاستراتيجي

النتيجة

توسّع مزيج منتجات التقييم وأصبحت التقييمات القطرية تهيمن عليه بشكل متزايد، وهذا مدفوع بشكل كبير بتوقعات المجلس وطلب الإدارة. وبينما تحظى تقييمات الاستراتيجيات القطرية والبرامج القطرية بالتقدير، فإن أصحاب المصلحة يعربون عن قلقهم من أن المزيج الحالي يرهق الموارد ولا يترك مساحة كافية للتعلم الاستراتيجي المواضيعي والمؤسسي.

الرسالة الرئيسية

لا تكمن المسألة الأساسية في الجودة بل في الانتقائية: فمن شأن وجود مزيج أكثر تركيزا من المنتجات، مع عدد أقل من التقييمات المختارة بشكل استراتيجي أكبر، حول مواضيع شاملة ذات أولوية للمؤسسة، أن يعزز بشكل كبير الملاءمة والأثر. كما يمكن لمكتب التقييم المستقل استخدام نتائج التقييم التي توصل إليها لتقديم موجز تجميعي للبلدان وملاحظات فورية تصدر في الوقت المناسب قبل المناقشات الرئيسية الرفيعة المستوى.

توصيات مكتب التقييم المستقل في الصندوق والمتابعة

النتيجة

تُعتبر توصيات مكتب التقييم المستقل المتعلقة بالتقييم عموماً مفيدة وملائمة، لكنها غالباً ما تكون مجزأة إلى توصيات فرعية، أو لا تُمنح الأولوية الكافية، أو تكون توجيهية بشكل مفرط، مما يجعل التنفيذ ومتابعة التنفيذ أمراً صعباً نظراً للقيود على الموارد والقدرة الاستيعابية لدى الصندوق.

الرسالة الرئيسية

من شأن تقليص عدد التوصيات وجعلها أكثر دقة واستراتيجية وتحديدًا للأولويات أن يعزز التنفيذ والتعلم.

بيئة التعلم والقدرة الاستيعابية

النتيجة

يتمتع الصندوق بثقافة تقييم قوية، ويحظى التعلم بالتقدير من حيث المبدأ. ولكن في الواقع، تُعد القدرة الاستيعابية للتعلم محدودة بسبب ضيق وقت موظفي الصندوق والقدرات التقييمية. وتؤدي منتجات التقييم المؤجلة والوصول إلى رؤى التقييم التي تتطلب وقتاً وجهداً إلى الحد من فرص التعلم وتصحيح المسار في الوقت المناسب.

الرسالة الرئيسية

يعتمد تعزيز التعلم بدرجة أقل على إنتاج المزيد من التقييمات، مقارنةً باعتماده على تقديم مخرجات استراتيجية، ومناسبة التوقيت، وموجزة وموجهة تتواءم مع دورات صنع القرار ومزيج أكثر تخصيصاً من منتجات التقييم، مما يتيح الاستفادة من الدروس وتطبيقها. كما تُعد سهولة الوصول إلى التقارير عاملاً أساسياً لخلق طلب على رؤى التقييم. ويمكن أن تكون الأدوات الجديدة المدعومة بالذكاء الاصطناعي نقطة تحول في توفير وصول سهل وهادف إلى نتائج ودروس التقييم من جانب العرض. ومع ذلك، يجب أيضاً تعزيز جانب الطلب لضمان أن يُقدّر التعلم ويُحفز من قبل المؤسسة.

الرسالة العامة لاستعراض الأقران الخارجي

- تتسم وظيفة التقييم في الصندوق بالمصداقية والاستقلالية وتحظى بتقدير كبير.
- بالنسبة لمكتب التقييم المستقل، يتمثل التحدي الرئيسي في التركيز الاستراتيجي والاستخدام، وليس الجودة أو الاستقلالية.
- تكتسي زيادة الانتقائية، وتحسين التكامل مع التقييم الذاتي، ووضع توصيات أكثر قابلية للتنفيذ أهمية بالغة لتعزيز الأثر.
- بالنسبة للتقييم الذاتي، يتمثل التحدي الرئيسي في ثبات الجودة وكذلك الملاءمة والفائدة لأغراض التعلم.
- تُعد معالجة قدرة التقييم (على المستوى القطري والمؤسسي) وقدرة المنظمة على الاستفادة من دروس التقييم عاملين محوريين لضمان استمرار التقييم في دعم الفعالية الإنمائية والكفاءة المؤسسية للصندوق.
- من شأن وضع استراتيجية للذكاء الاصطناعي أن يساعد على تحديد فرص زيادة الكفاءة والفعالية وتوفير إطار لاستخدام مستدام وأخلاقي للذكاء الاصطناعي في التقييم. وينبغي أن يتم ذلك، كما هو الحال في تنمية القدرات، بالتعاون الوثيق والاعتماد المتبادل مع المؤسسات المالية الدولية الأخرى والشركاء الآخرين.

1- المقدمة والخلفية

1-1 المقدمة

- 1- يبدأ استعراض الأقران الخارجي وفقا لمتطلبات سياسة التقييم في الصندوق ويهدف إلى تقديم تقييم خارجي لوظيفة التقييم في الصندوق، بما في ذلك التقييم الذاتي الذي تقوده الإدارة والتقييم المستقل من قبل مكتب التقييم المستقل.
- 2- ويتمثل الغرض الرئيسي من استعراض الأقران الخارجي في تقييم وظيفة التقييم في الصندوق لتعزيز مساهمته في الفعالية الإنمائية والكفاءة المؤسسية للصندوق.
- 3- ويأتي هذا الاستعراض في مرحلة مفصلية حرجة، في سياق البيئة التشغيلية الحالية التي يواجهها الصندوق. فقد أدى عقد من الإصلاحات المؤسسية وتطور التوقعات المتعلقة بالفعالية الإنمائية إلى تحويل المنظمة لتكون أفضل تجهيزا لتلبية الاحتياجات المتغيرة للأمن الغذائي والإنتاج الزراعي في أفقر مناطق البلدان في سياق عالمي يشهد حالة من عدم اليقين البيئي والمالي والسياسي. وقد وضع الصندوق هيكلياً مالية جديدة يركز بشكل أكبر على عمليات القطاع الخاص والوصول إلى مصادر تمويل بديلة بهدف زيادة الفعالية التشغيلية والأثر الميداني. ومن العوامل السياقية المهمة الدورة القادمة لتجديد موارد الصندوق (التجديد الرابع عشر لموارد الصندوق للفترة 2028-2031). كما يشكل اتجاه تقليص المساعدات الإنمائية الرسمية مصدر قلق متزايد في الصندوق.
- 4- ومن الأمور ذات الصلة بشكل خاص بوظيفة التقييم إعادة معايرة مكتب الفعالية الإنمائية باعتباره مركزاً مؤسسياً للأدلة والفعالية، بوظائف تتعلق بضمان الجودة، والنتائج، والموارد، وإدارة المعرفة، والابتكار، وتقييم الأثر. وفي عام 2026، أضيفت وظيفة استراتيجية مؤسسية إلى مكتب الفعالية الإنمائية لتعزيز الجهود الاستراتيجية في المؤسسة والربط فيما بينها بشكل أفضل، وذلك بهدف دعم وتفعيل الرؤية الاستراتيجية التي وضعها الرئيس. ويهدف ذلك إلى ضمان الاتساق الاستراتيجي بين الأولويات المؤسسية، والاتجاهات المالية والتشغيلية، والأدلة، والتموضع المؤسسي، وأولويات المشاركة الخارجية.
- 5- وفي حين يتسم الاستعراض بطابع استرجاعي، حيث يقيم أداء وظيفة التقييم منذ تقرير استعراض الأقران الخارجي لعام 2019 والالتزام بالمعايير المتفق عليها من قبل فريق التعاون في مجال التقييم التابع للمصارف الإنمائية المتعددة الأطراف، وفريق الأمم المتحدة المعني بالتقييم، وشبكة التقييم التابعة للجنة المساعدة الإنمائية التابعة لمنظمة التعاون والتنمية في الميدان الاقتصادي، فإنه تكويني واستشراقي بشكل مقصود، ويتوقع أيضاً ما هو مطلوب لتعزيز وظيفة التقييم في المستقبل القريب.
- 6- واتخذ الفريق وجهة نظر محايدة، استناداً إلى المعرفة بأفضل الممارسات، وفيما يتعلق بالمعايير الأخلاقية والمعلومات السرية. ويتوافق نهج الاستعراض ومنهجيته مع النهج والمنهجية المتفق عليهما من قبل فريق التعاون في مجال التقييم وفريق الأمم المتحدة المعني بالتقييم وشبكة التقييم. وعلى النحو المبين في إطار التقييم الخاص باستعراض الأقران الخارجي (الذيل الأول) الذي يستجيب لورق النهج بشأن استعراض الأقران الخارجي، يبنى التقييم على مصادر متعددة للمعلومات بما في ذلك عمليات الاستعراض المنهجية للوثائق، والمقابلات، وجلسات مجموعات التركيز، واستقصاء لجميع الموظفين وأعضاء المجلس التنفيذي (انظر الذيل الثاني لمزيد من التفاصيل بشأن المنهجية).
- 7- وبما يتماشى مع أفضل الممارسات في عمليات الاستعراض الخارجية لوظائف التقييم، قامت الإدارة (مكتب الفعالية الإنمائية) ومكتب التقييم المستقل بإجراء تقييمات ذاتية للتقييم الذاتي والتقييم المستقل على التوالي. وتوفر تقارير

التقييمات الذاتية هذه معلومات قيمة لاستعراض الأقران الخارجي. وعلى الرغم من أن التقييمات الذاتية لا تخضع للتحقق من قبل فريق استعراض الأقران الخارجي، فهي توفر معلومات قيمة لقراء هذا التقرير. وتقارير التقييمات الذاتية متاحة عند الطلب من خلال الاتصال بالإدارات المعنية.

8- وعلى عكس النسختين السابقتين من استعراض الأقران الخارجي (2009 و2019)، يشمل هذا الاستعراض كلا من التقييم الذاتي الذي تقوده الإدارة والتقييم المستقل الذي يجريه مكتب التقييم المستقل، كما أنه يقيم مدى مواءمة وتكامل جزئي وظيفة التقييم.

2-1 الخلفية – تنفيذ توصيات استعراض الأقران الخارجي لعام 2019

9- تشكل توصيات استعراض الأقران الخارجي لعام 2019 الخلفية لاستعراض الأقران الخارجي الحالي. وقد قدمت التوصيات الثماني في سياق الصندوق في ذلك الوقت. وتظهر نتائج استعراض الأقران الخارجي هذا أن جميع التوصيات قد نُفذت.

10- ووفقا للتقييمات الذاتية لمكتب التقييم المستقل والإدارة، كان التنفيذ واسع النطاق حيث تضمن تعديلات على سياسات التقييم، وتعزيز الإرشادات المنهجية، وتحسين عمليات الإشراف والإنجاز، وتطبيق إطار الفعالية الإنمائية، وإنشاء مكتب الفعالية الإنمائية. ووفقا للتقييم الذاتي للإدارة، تتطلب المرحلة التالية دمج معايير جديدة في الممارسة اليومية من خلال الحوافز، وبناء القدرات، وتدفقات العمل الواضحة.

11- وترد أدناه بعض تفاصيل التنفيذ:

12- **تعديل سياسة التقييم.** نُشرت سياسة التقييم المُعدلة في الصندوق في عام 2021. ولأول مرة، شملت السياسة كلا من التقييم المستقل والتقييم الذاتي، مما وسع نطاق وشمولية ممارسات التقييم داخل الصندوق.

13- **تبسيط العمليات والإجراءات الداخلية لمكتب التقييم المستقل.** أظهرت عمليات الاستعراض الحديثة تحسنا في كفاءة تقييم الاستراتيجية القطرية والبرنامج القطري. وجرى تقليص الجدول الزمني لتقييمات الاستراتيجيات القطرية والبرامج القطرية بنحو 10 في المائة، حيث غالبا ما تُنجز التقييمات قبل الموعد المحدد، مما يعكس تحسنا في التخطيط والتنفيذ. وقد تمكّن مكتب التقييم المستقل من مواكبة التوسع في تغطية هذه التقييمات بزيادة نقل عن 30 في المائة في الميزانية وموارد الموظفين. وانخفض متوسط الميزانيات المقررة والفعالية لهذه التقييمات (بحوالي 9 في المائة و22 في المائة على التوالي)، كما ساعدت حلقات العمل الافتراضية على خفض التكاليف بشكل أكبر.

14- **تعديل مزيج منتجات مكتب التقييم المستقل.** قام مكتب التقييم المستقل بتوزيع مخرجاته، بما يشمل التقييمات المواضيعية والمؤسسية، وسلاسل الاستعراض، ومذكرات التعلم، والأوراق البحثية، والكتب، والمقالات الخاضعة لاستعراض الأقران. وقُدمت منتجات جديدة مثل المخططات البيانية، وسلاسل الفيديو، والبودكاست، والمدونات، ومجلة مستقلة لتعزيز التواصل والمشاركة.

15- **إعداد استراتيجية متعددة السنوات لمكتب التقييم المستقل من أجل تنفيذ سياسة التقييم.** تحدد استراتيجية التقييم المتعددة السنوات (2022-2027) أربعة أهداف رئيسية، منها تعزيز ثقافة مؤسسية شفافة وموجهة نحو التعلم ومسؤولة، والمشاركة مع الإدارة والدول الأعضاء والشركاء الخارجيين. ووفقا لمكتب التقييم المستقل¹، توفر الاستراتيجية إطارا لزيادة تغطية التقييمات القطرية خلال العامين الماضيين بزيادة بنسبة 333 في المائة في

¹ التقييم الذاتي لمكتب التقييم المستقل في الصندوق.

التغطية القطرية وزيادة بنسبة 1 578 في المائة في القيمة بالدولار للحفاظ الخاضعة للتقييم، على أن يُدار كل ذلك بزيادة أقل من 30 في المائة في الميزانية والموظفين.

16- **إعداد دليل جديد للتقييم.** في عام 2022، جرى توقيع الإصدار الثالث من دليل التقييم في الصندوق بالتعاون مع الإدارة - وهي سابقة أولى من نوعها في أي وكالة من وكالات الأمم المتحدة أو مؤسسة مالية دولية. ويشمل هذا الدليل عمليات التقييم المستقل والتقييم الذاتي على حد سواء، ويعكس نهجا أكثر تكاملا في التقييم داخل الصندوق.

17- **تبسيط وتنسيق ميزانية مكتب التقييم المستقل مع الميزانية الإدارية الإجمالية للصندوق.** على الرغم من التوسع الكبير في تغطية التقييم والأنشطة خلال العامين الماضيين، حافظ مكتب التقييم المستقل على زيادات الميزانية عند أقل من 30 في المائة. وانخفض متوسط التكلفة لكل تقييم للاستراتيجية القطرية والبرنامج القطري، وخُفض الإنفاق الإجمالي من خلال تبسيط العمليات واعتماد الأدوات الرقمية.

النقاط الرئيسية

يُجرى استعراض الأقران الخارجي لعام 2026 وفقا لسياسة التقييم في الصندوق وبما يتماشى مع المعايير الدولية. ويعتمد الاستعراض على مصادر بيانات متعددة ومشاورات موسعة. وهو رفيع المستوى واستشراقي بشكل مقصود.

وأدى استعراض الأقران الخارجي لعام 2019 إلى إصلاحات مهمة (مراجعة السياسات، ومواءمة التصنيفات، وإنشاء مكتب الفعالية الإنمائية). ويخلص الاستعراض الحالي إلى أن الاتجاه صحيح، غير أن عمق التنفيذ وتحديد الأولويات والمرونة، بالإضافة إلى الاتساق على مستوى المنظومة ككل، أصبحت أكثر أهمية من مواصلة التوسع.

2- النتائج الرئيسية

1-2 التقييم العام: وظيفة قوية يتعين أن تكون أكثر استراتيجية لمواكبة الصندوق في المستقبل

- 18- يخلص استعراض الأقران الخارجي إلى أن وظيفة التقييم (بما في ذلك التقييم المستقل والتقييم الذاتي على حد سواء) في السياق الحالي تؤدي أداء جيداً دون أي أوجه قصور جوهرية. وكما هو موضح أعلاه، فقد نُفذت توصيات استعراض الأقران الخارجي لعام 2019. وتحظى الوظيفة بشكلها الحالي بالتقدير والاحترام على نطاق واسع في الصندوق.
- 19- وتتسق سياسة التقييم ودليل التقييم مع الممارسات الشائعة في المؤسسات المالية الدولية ومنظمات الأمم المتحدة الأخرى. ويخلص فريق استعراض الأقران الخارجي إلى أن السياسة واضحة ومرنة بما يكفي لتشكيل إطاراً لمزيد من التحسينات في وظيفة التقييم. وينبغي توضيح بعض العناصر في الملاحق (مثل أدوار ومسؤوليات مكتب الفعالية الإنمائية، ووضوح العمليات والممارسات المتعلقة بالموارد البشرية)، ولكن لا حاجة إلى تنقيح السياسة نفسها.
- 20- ووصل التقييم المستقل الذي يديره مكتب التقييم المستقل إلى مرحلة النضج² خلال السنوات القليلة الماضية مع التنفيذ الفعال لمزيج من منتجات التقييم ونشر النتائج والتوصيات والتواصل بشأنها على نطاق واسع. وتُستوفى إلى حد كبير المعايير الأساسية لوظيفة تقييم مستقلة موثوقة، فيما يتعلق بالمصادقية والاستقلالية والفائدة.
- 21- والتقييم الذاتي، رغم تاريخه الطويل في الصندوق، هو في مرحلة مبكرة من النضج مقارنة بوظائف مماثلة في مؤسسات مالية دولية أخرى؛ ومع ذلك، فقد حدد نقاط ضعفه بكل مصداقية وأسس هيكلًا متقناً وخططاً مقنعة للتطبيق (راجع التقييم الذاتي لمكتب التقييم المستقل). وتسهم إعادة معايرة مكتب الفعالية الإنمائية وتعيين مدير إداري للفعالية الإنمائية والاستراتيجية المؤسسية في خلق توقعات إيجابية على نطاق المؤسسة.
- 22- وعلى الرغم من هذا التقييم الإيجابي عموماً، فمن غير المؤكد مدى استعداد وظيفة التقييم لمواكبة المستقبل المتوقع. وبالنظر إلى أن الصندوق يقوم بإصلاحات مقصودة بهدف الاستجابة الفعالة لسياق عالمي متزايد التحديات، يظل من غير الواضح ما إذا كانت وظيفة التقييم متوائمة بما فيه الكفاية مع هذه الإصلاحات لتتمكن من أداء دورها المنشود في المستقبل. ولا يزال أمام مكتب الفعالية الإنمائية طريق طويل قبل التنفيذ الكامل لهيكل التقييم الذاتي وقبل أن تصبح جودة التقييم الذاتي مقبولة على نطاق المنظمة ككل. ولم يُكَيَّف حالياً مزيج منتجات مكتب التقييم المستقل وقدرته بشكل كامل لمواكبة المستقبل كما هو موضح في جدول أعمال الإصلاح في الصندوق.
- 23- وفي بيئة سريعة التغير، لا يمكن لوظيفة التقييم (سواء التقييم المستقل أو التقييم الذاتي) أداء دورها بالاعتماد في المقام الأول على توفير أدلة بشأن أداء العمليات المُنجزة. ويجب إيلاء مزيد من الاهتمام للعمليات المقررة والجارية واهتمام أقل للامتثال والعمليات السابقة. ويلزم أن يتحول التركيز من الطابع الاسترجاعي إلى الطابع الاستشراقي. وتُعدّ الجودة عند الإدراج مهمة لتحقيق النتائج في المستقبل. ووظيفة التقييم (ولا سيما مكتب التقييم المستقل) لديها القدرة على تقديم أدلة ذات أهمية استراتيجية أكبر³ للإدارة العليا والمجلس التنفيذي. ويتطلب ذلك الانتقائية والتركيز وحسن التوقيت والكفاءة. ويجب أن تكون خطط التقييم لدى مكتب التقييم المستقل متوائمة بشكل وثيق مع عمليات تجديد الموارد ونقاط القرارات الاستراتيجية المؤسسية الرئيسية الأخرى.

² مقارنة بالوضع الموصوف في استعراض الأقران الخارجي لعام 2019.

³ "استراتيجية" تعني أن تكون استباقية من خلال تقديم أدلة في الوقت المناسب بشأن المواضيع الحالية وإيلاء الاهتمام للمسائل الهامة في المستقبل

24- ويُعدّ الاستخدام المتزايد للذكاء الاصطناعي من العناصر الأساسية الأخرى في وظيفة التقييم الاستباقية. فالذكاء الاصطناعي يوفر فرصا لتحسين تغطية التقييم، وتوقيته، واستخدام الأدلة. ولكنه يجلب في نفس الوقت تحديات تتعلق بالأخلاقيات والتحيز والمساءلة والمهارات والحوكمة، وهي تحديات تحتاج إلى إدارة متأنية. وتشمل بعض هذه الفرص الحصول على توليف أسرع للأدلة والتقييم الجامع، والوصول إلى مصادر بيانات أوسع ومعرفة تقييمية، واكتساب الكفاءة في عمليات التقييم. وفي الوقت نفسه، من المهم جدا الإقرار بالتحديات المتعلقة بالأخلاقيات وحقوق الإنسان والعدالة والتحيز والتمثيل والمساءلة والشفافية. وتجدر الإشارة إلى مبادئ الذكاء الاصطناعي في منظمة التعاون والتنمية في الميدان الاقتصادي⁴ التي تدعو إلى إمكانية تتبع مجموعات البيانات والعمليات والقرارات. ويعني ذلك توثيق مدى تأثير أدوات الذكاء الاصطناعي على النتائج وضمان أن تظل الأحكام التقييمية تُنسب بوضوح إلى المقيمين من البشر.

25- وكما هو موضح في إطار تقييم استعراض الأقران الخارجي (الذيل الأول)، يركز الاستعراض على ستة معايير أساسية: المصادقية، والاستقلالية، والفائدة، والتكامل بين التقييم الذاتي والتقييم المستقل، والإدارة والموارد البشرية، وتهيئة بيئة تمكينية للتعليم والتطور. وترد أدناه النتائج الرئيسية المتعلقة بهذه المعايير.⁵

النقاط الأساسية

تتسم وظيفة التقييم في الصندوق بالمصادقية والاستقلالية وتحظى بتقدير كبير، كما أنها تُقارن بشكل إيجابي مع نظيراتها من مؤسسات التمويل الدولية ووكالات الأمم المتحدة الأخرى.

ولكن لكي تظل هذه الوظيفة ملائمة ومؤثرة في سياق مؤسسي سريع التغير، خصوصا في ظل الهيكلة المالية الجديدة للصندوق والتجديد الرابع عشر القادم لموارد الصندوق وتنامي دور للذكاء الاصطناعي، يجب أن تصبح الوظيفة أكثر استراتيجية وانتقائية وتوجهها نحو التعلم بشكل ملحوظ، وذلك من خلال منتجات أكثر ملائمة ومناسبة التوقيت، ومتابعة أقوى، وتعزيز المواءمة بين التقييم المستقل والتقييم الذاتي.

ويُعدّ إيلاء الاهتمام للفرص والتحديات التي تنطوي عليها زيادة استخدام الذكاء الاصطناعي أمرا أساسيا.

2-2 المصادقية: قوية وموثوقة، مع الحاجة إلى اتباع نهج أكثر استراتيجية

26- تقيّم معايير المصادقية مدى مواءمة وظيفة التقييم مع الممارسات المقبولة عندما يتعلق الأمر بجودة العمل، بما في ذلك منهجية جمع وتحليل البيانات، وحسن التوقيت، والشفافية، ومشاركة أصحاب المصلحة في عمليات التقييم، وتطبيق المعايير الأخلاقية في عمليات العمل، وأهمية مسائل التقييم.

27- وتتبع الوثائق الإرشادية للتقييم معايير مشتركة. وتتواءم الوثائق الإرشادية الرئيسية للتقييم (سياسة التقييم لعام 2021 ودليل التقييم لعام 2022) مع معايير فريق التعاون في مجال التقييم وفريق الأمم المتحدة المعني بالتقييم وشبكة التقييم، وتتماشى مع وثائق مماثلة في منظمات أخرى متعددة الأطراف. ويُستثنى من ذلك أن إعادة معايرة مكتب الفعالية الإنمائية لم تنعكس بعد بالكامل في هذه الوثائق.

⁴ منظمة التعاون والتنمية في الميدان الاقتصادي. توصية مجلس الذكاء الاصطناعي، OECD/LEGAL/0449.
⁵ يمكن تقديم لمحة عامة مفصلة عن الإجابات على أسئلة التقييم الواردة في إطار تقييم استعراض الأقران الخارجي، عند الطلب.

28- ويُعدّ دور المساءلة أقوى من دور التعلم، بينما يجد المجيبون في المقابلات وفي الاستقصاء أن وظيفة التقييم تؤدي دور المساءلة ودور التعلم في آن واحد، فإن دور المساءلة مهيم في الوقت الحالي (الشكل 1).⁶

الشكل 1: سحابة الكلمات مع الإجابات الواردة في الاستقصاء على سؤال الكلمة المفتاحية المرتبطة بمصطلح "التقييم"



29- **يحظى مكتب التقييم المستقل بمصادقية عالية.** يؤكد الاستقصاء (الذيل الرابع) والمقابلات أن عمل مكتب التقييم المستقل يعتبر ذا مصداقية على نطاق المنظمة ككل. وعلّق أحد المقيّمين على الاستقصاء قائلا: "أدى مكتب التقييم المستقل دورا أساسيا في تعزيز مصداقية تدخلات الصندوق وفي حث الإدارة على التحسين عند الحاجة."

30- كما يجري تطوير ثقافة التقييم في الصندوق. وقد اطلع نحو نصف المجيبين على الاستقصاء على منتجات مكتب التقييم المستقل وأفاد 40 في المائة منهم بأنهم استخدموها لإرشاد عملهم. وأفاد 7 في المائة فقط من المجيبين تحديدا بأنهم لم يشاركوا قط في أي أعمال تقييم أجريت في الصندوق. وهذه النتيجة إيجابية مقارنة مثلا بالمصرف الأوروبي للإنشاء والتعمير.⁷

31- ولدى مكتب التقييم المستقل آليات راسخة لضمان الجودة - مع بعض الاستثناءات - داخليا وخارجيا على حد سواء مع فريق استشاري للتقييم.⁸ فعلى سبيل المثال، لا يزال تطبيق المعايير الأخلاقية على المستوى العملي غير مكتمل، رغم أن موظفي مكتب التقييم المستقل يجدون أن الالتزام بالمعايير الأخلاقية قد تحسن في الآونة الأخيرة، مثل إدراج الاعتبارات الأخلاقية في جميع أوراق النهج وإعداد إرشادات داخلية حول جمع البيانات الأخلاقية في المجتمعات المحلية. وقد تلزم مواصلة تعزيز القدرات الداخلية من أجل التطبيق المتسق للمعايير.

32- **زيادة الكفاءة في مكتب التقييم المستقل في الصندوق.** وفقا للتقييم الذاتي لمكتب التقييم المستقل، فقد تحسنت الكفاءة في تقديم تقييمات الاستراتيجيات القطرية والبرامج القطرية من 10.38 شهر في 2019 إلى 9.29 شهر في عامي 2024/2025. وانخفض متوسط ميزانية تقييم الاستراتيجية القطرية والبرنامج القطري لتكاليف غير الموظفين

⁶ يعلق المجيبون على التركيز على المساءلة، بما في ذلك ما يلي: "يبدو أن مكتب التقييم المستقل بضطلع بدور أقرب إلى الامتثال منه إلى دعم التعلم مع الأقران." "ويمكن نشر الدروس المستفادة من تقارير إنجاز المشروعات بشكل أفضل لإرشاد الجيل الجديد من المشروعات في البلد."

7. خلص استقصاء التصورات لعام 2025 الذي يجريه المصرف الأوروبي للإنشاء والتعمير إلى أن الوعي والمعرفة بدائرة التقييم المستقل كانتا متفاوتتين على نطاق المؤسسة. وبينما أبلغ 100 في المائة من أعضاء المجلس والإدارة العليا عن معرفتهم بدائرة التقييم المستقل، لم يبلغ عن ذلك من الموظفين سوى 31 في المائة.

⁸ يتألف الفريق الاستشاري للتقييم من خبراء دوليين مختارين في التقييم.

في نفس الفترة من 202 301 دولار أمريكي إلى 156 483 دولارا أمريكيا، بينما لا توجد مؤشرات على تراجع الجودة.

33- نتائج متباينة فيما يتعلق بمشاركة مكتب التقييم المستقل مع أصحاب المصلحة. ويتبع مكتب التقييم المستقل، من حيث المبدأ، نهجا استشاريا في جمع البيانات، معتمدا كذلك على بيانات البلدان. ومع ذلك، يتمثل أحد التحديات المتكررة في القدرة القطرية على جمع بيانات موثوقة عن المشروعات. ويشير موظفو مكتب التقييم المستقل إلى إمكانية تعزيز الشمولية ومشاركة المستفيدين بدرجة أكبر في مراحل أخرى من عملية التقييم، مع الحفاظ على الاستقلالية. وأجاب 21 في المائة من المجيبين على الاستقصاء بالجواب "أوافق بشدة" على البيان بأن مكتب التقييم المستقل يشارك بشكل كاف مع أصحاب المصلحة في الصندوق، بينما أجاب 49 في المائة بالجواب "أوافق" وأجاب 12 في المائة بالجواب "لا أوافق". وقدم أحد المجيبين على الاستقصاء التوصية التالية: "أعتقد أن المزيد من المشاركة مع المستفيدين سيحسن عمل الصندوق وسيقلل من مخاطر إساءة الاستخدام من جانب وحدات إدارة المشروعات والمسؤولين الحكوميين".

34- ويقر مكتب التقييم المستقل والمجيبون الآخرون بالقدرة القطرية المحدودة على الرصد والتقييم. وهذا أحد العوامل المقيدة لموثوقية ومصداقية بيانات المشروعات. وبدلا من قيام مكتب التقييم المستقل بنفسه ببناء قدرات التقييم على المستوى القطري، فقد قرر التعاون مع مبادرة التقييم العالمية، بما في ذلك مراكز التعلم الإقليمية الخاصة بالتقييم والنتائج. ويدعم فريق استعراض الأقران الخارجي هذا النهج.

35- ويختلف مزيج المنتجات في برنامج عمل مكتب التقييم المستقل الذي تبلغ مدته ثلاث سنوات عن برامج العمل الخاصة بالمؤسسات المالية الدولية الأخرى. ويتمثل الاختلاف الرئيسي في النسبة الكبيرة من تقييمات الاستراتيجيات القطرية والبرامج القطرية بين التقييمات المنشورة البالغ عددها 40 تقييما في فترة الاستعراض (2019-2025). وأنجز مكتب التقييم المستقل 29 تقييما للاستراتيجيات القطرية والبرامج القطرية، و5 تقييمات مؤسسية، وتقييمين من التقييمات المواضيعية، وتقييما واحدا لمجموعة مشروعات، وتقييمين من التقييمات دون الإقليمية (الشكل 2).

الشكل 2: تقييمات مكتب التقييم المستقل المنشورة للفترة 2019-2025



36- ونأتي التغطية الواسعة لتقييمات الاستراتيجيات القطرية والبرامج القطرية نتيجة لتزايد الطلب من العمليات والمجلس التنفيذي، وهي تملأ الثغرات السابقة في التقييمات الموثوقة لتنفيذ برامج الفرص الاستراتيجية القطرية في الصندوق. ومن بين المجيبين على الاستقصاء، قام 39 في المائة بتصنيف تقييمات الاستراتيجيات القطرية

والبرامج القطرية كأكثر منتجات مكتب التقييم المستقل فائدة، بينما صنف 36 في المائة التقييمات المواضيعية بأنها الأكثر فائدة، ورأى 34 في المائة أن التقييمات المؤسسية كانت هي الأكثر فائدة.

37- **جودة تقييمات مكتب التقييم المستقل جيدة عموماً لكنها قابلة لمزيد من التحسين.** أعرب معظم المجيبين على الاستقصاء عن موافقتهم (منهم 29 في المائة يوافقون بشدة) على البيان بأن نهج ومنهجيات مكتب التقييم المستقل صارمة.

38- وفي إطار عملية استعراض الأقران الخارجي، أُجري استعراض لعينة من 20 تقييماً لمكتب التقييم المستقل لتقديم لمحة عامة على المنهجيات والجودة العامة لمنتجات التقييم (انظر الذيل الخامس). وتشمل العينة جميع التقييمات المؤسسية والمواضيعية وتقييمات مجموعات المشروعات والتقييمات دون الإقليمية التي نُشرت خلال فترة الاستعراض، بالإضافة إلى تسعة تقييمات للاستراتيجيات القطرية والبرامج القطرية، جرى اختيارها عشوائياً. وبشكل عام، تظهر التقييمات الخاضعة للاستعراض جودة عالية وتعكس تقديراً قوياً للأساليب والأدلة، حتى في سياق نطاقات التقييم الطموحة ومواد التقييم المعقدة.

39- **وفي الوقت نفسه، جرى تحديد عدة مجالات للتحسين.** وتشمل هذه المجالات زيادة الشفافية فيما يتعلق بجودة البيانات والتتليث، وإيلاء اهتمام أكثر منهجية للبيئة التي يعمل فيها الصندوق، وتعميق النظر في الصلاحية الخارجية للنتائج.⁹ ولا توفر معظم التقييمات وضوحاً كافياً بشأن كيفية تثليث الأدلة المستقاة من مصادر مختلفة واستخدامها لدعم النتائج المتعلقة بأسئلة التقييم الأولية. كما وجد الاستعراض تركيزاً محدوداً على الفهم السياقي، وهو ما قد يقيد الفهم الأعمق للبيانات التي تعمل فيها التدخلات المدعومة من الصندوق. وسيعمل تعزيز استخدام نظرية التغيير، ولا سيما صياغتها فيما يتعلق بالتحليل التجريبي، على تحسين الوضوح التحليلي بشكل أكبر. وعلى نطاق أوسع، فإن اعتماد نهج أكثر عقلانية ورسمية في تصميم التقييم، يتجاوز مجرد تجميع الأساليب والأدوات نحو إطار متسق ومنظم، من شأنه أن يدعم فهماً أعمق لجودة التقييم. كما يمكن لدمج النهج المبتكرة، مثل تحليل المساهمات، أن يعزز تقييمات الاتساق والاستدلال السببي.

40- ومن شأن توسيع الاهتمام بهذه الأبعاد أن يحسّن صحة التقييمات المستقبلية وموثوقيتها وفائدتها العامة، مع الحفاظ على المعايير العالية للدقة المنهجية الواضحة بالفعل في العينة الخاضعة للتحليل.

41- **وسيساهم وضع أولويات أكثر استراتيجية في برامج عمل مكتب التقييم المستقل في تعزيز الاتساق والأثر.** وتظهر الوثائق المقدمة إلى مجلس محافظي الصندوق ومدخلات مكتب التقييم المستقل أن برنامج عمل مكتب التقييم المستقل يتبع ممارسات جيدة شائعة بين المؤسسات المالية الدولية فيما يتعلق بالتشاور مع الإدارة والهيئات الرئاسية في الصندوق. وتمنح المشاورات الإدارة والمجلس إمكانيات لطرح المجالات ذات الأولوية، وإبراز برامج الفرص الاستراتيجية القطرية ومسائل التخرج، ومناقشة مجالات ذات أهمية استراتيجية أو شاملة على المستوى المؤسسي. وتسترشد التقييمات على المستويات الأعلى بمدخلات الإدارة، وتغطية التقييمات، والأهمية الاستراتيجية، والمسائل الجوهرية التي طُرحت في التقييمات السابقة، وتوقيت تنفيذ الاستراتيجيات و/أو السياسات وتعديلاتها. ويُعدّ مكتب التقييم المستقل مذكرات أولية لتحديد النطاق تلقي مزيد من الضوء على المجالات الرئيسية واعتبارات التوقيت. فعلى سبيل المثال، قرر مكتب التقييم المستقل في عام 2026 تحويل استعراض سريع لمسائل الموارد البشرية تقييم مؤسسي كامل، مشيراً إلى ضعف تغطية التقييم ونقص المسائل الاستراتيجية التي طُرحت في تقييمات اللامركزية وتقييمات التجديدين الحادي عشر والثاني عشر لموارد الصندوق. كما جرى إعطاء أولوية

⁹ تشير الصلاحية الخارجية (أو القابلية للتعميم) إلى مدى إمكانية تطبيق نتائج أي تقييم في بيئات أخرى.

للمجالات الناشئة ذات الصلة بتحديد الاتجاهات المستقبلية للصندوق، بما في ذلك التمويل التكميلي ومشاركة القطاع الخاص.

42- وعلى وجه الخصوص، قدمت التقييمات المؤسسية رؤى بشأن مواضيع استراتيجية مهمة. ومن الأمثلة على ذلك تقييم اللامركزية، وتقييم التجديدين الحادي عشر والثاني عشر لموارد الصندوق. وعُلق أحد المقيمين على الاستقصاء قائلاً: "كان تقييم التجديدين الحادي عشر والثاني عشر لموارد الصندوق مفيداً بشكل خاص لأنه أثار صراحة العديد من المسائل التي لا تزال ذات صلة بالمنظمة بأكملها". وأدلى العديد من المشاركين في المقابلات بتعليقات مماثلة.

43- وعلى الرغم من المشاورات الواسعة والاعتبارات الداخلية، تبرز شكوك بشأن الأهمية الاستراتيجية لبعض أجزاء برنامج العمل، ولا سيما مواصلة إجراء تقييمات المشروعات¹⁰ وبدء تقييمات الاستراتيجيات القطرية والبرامج القطرية عندما تكون حافظة الصندوق صغيرة جداً أو عندما لا يكون برنامج الفرص الاستراتيجية القطرية مخططاً له. ويتمتع مكتب التقييم المستقل بالقدرة على التعمق في التفاصيل. كما أن الطبيعة المدفوعة بالطلب التي يتسم بها برنامج العمل - والمدفوعة بطلبات الإدارة والمجلس فيما يتعلق بالتقييمات القطرية - قد تُبعد مكتب التقييم المستقل عن الاضطلاع بدور أكثر تأثيراً. وقد أشار العديد من أعضاء المجلس إلى تقييم التجديدين الحادي عشر والثاني عشر لموارد الصندوق¹¹ باعتباره نموذجاً مثالياً للتقييمات الاستراتيجية ذات القيمة العالية. وتوصف هذه التقييمات بأنها "قيمة وتحظى بتقدير كبير"، ويرجع ذلك تحديداً إلى أنها أثارت مسائل مؤسسية صعبة بطريقة صريحة ومفيدة.

44- والحجة هي أن هذه القدرة ينبغي أن تركز على أهم المسائل التي يمكن تقديم معرفة جديدة فيها أو على مواضيع تنصدر جدول أعمال المجلس والإدارة العليا. ويعني ذلك أن إجراء تقييمات المشروعات وتقييمات الاستراتيجيات القطرية والبرامج القطرية الكاملة في بعض الحالات أقل فائدة بالنسبة لمكتب التقييم المستقل. ومن الأمثلة على ذلك تقييمات الاستراتيجيات القطرية والبرامج القطرية في بيرو، حيث يبدو أنه لم يُنجز سوى مشروع واحد منذ تقييم الاستراتيجية القطرية والبرنامج القطري السابق. ويؤكد فريق استعراض الأقران الخارجي أنه ينبغي دائماً وجود ارتباط وثيق بين تقييم الاستراتيجية القطرية والبرنامج القطرية وأي برنامج جديد للفرص الاستراتيجية القطرية. ومن الشواغل الأخرى أن تقييمات الاستراتيجيات القطرية والبرامج القطرية تستخدم نهجاً معيارياً. فأحياناً يكون النطاق واسعاً للغاية (مثل تغطية عشر سنوات في بلد شهد تغييرات كبيرة) أو لا يعالج التحديات الحالية بشكل كافٍ (مثل عدم مناقشة التخرج في بلد تُطرح فيه هذه المسألة على جدول الأعمال). كما يمكن أن تساعد زيادة التركيز والانتقائية في النطاق على ضمان استخدام قدرات مكتب التقييم المستقل في المجالات الأكثر فائدة. وسيساعد ذلك على إطلاق العنان لإمكانات مكتب التقييم المستقل لكي يضطلع بدور استباقي واستراتيجي أكثر أهمية في المستقبل. ويُعدّ الاستخدام الواعي والمدار جيداً للذكاء الاصطناعي مهما في هذا السياق.

45- وينبغي أن يعمل مكتب التقييم المستقل على إعادة توازن برنامج عمله لإعطاء الأولوية لما يلي:

- التقييمات المؤسسية الاستراتيجية التي تتواءم مع دورات تجديد الموارد وعمليات الإصلاح المؤسسي الكبرى؛

¹⁰ يشير مكتب التقييم المستقل إلى أن عدد تقييمات أداء المشروعات التي أُجريت قد انخفض بشكل كبير، من 10-12، إلى 4-5 كحد أقصى عندما لا تزال المشروعات جارية

¹¹ التقييم المؤسسي للأداء المؤسسي والتشغيلي في إطار التجديدين الحادي عشر والثاني عشر لموارد الصندوق. الوثيقة: EC 2026/132/W.P.7

- التقييمات المواضيعية والشاملة بشأن مواضيع ذات صلة مؤسسية مباشرة (الأمن الغذائي، ومشاركة القطاع الخاص، واستدامة الأثر، والذكاء الاصطناعي)؛
- التقييمات الجامعة و/أو الدراسات التوليفية المستندة إلى النتائج الحالية لتقييم الاستراتيجية القطرية والبرنامج القطري، بدلا من بدء تقييمات كاملة للاستراتيجيات القطرية والبرامج القطرية على حوافز محدودة أو بدون برنامج فرص استراتيجية قطرية لاحق.
- 46- وكإرشاد لتحديد الأهمية الاستراتيجية لأي تقييم، تجدر الإشارة إلى المشورة التي قدمها الفريق الاستشاري للتقييم بشأن ما يشكل تقييمات مؤثرة: (1) التوجه الاستراتيجي والملاءمة السياسية؛ (2) المنهجية الصارمة والمصدقية؛ (3) مشاركة أصحاب المصلحة؛ (4) وضوح النتائج؛ (5) قابلية تنفيذ التوصيات؛ (6) الأفق الزمني للمعرفة التقييمية؛ (7) حسن التوقيت؛ (8) النشر والتواصل بشأنها.
- 47- **تشهد مصداقية التقييم الذاتي تطورا.** ويمثل إطار الفعالية الإنمائية العمود الفقري لهيكلية الأداء والنتائج في الصندوق. ودمج إطار الفعالية الإنمائية المحدث 2.0 لعام 2021 تدفقات مختلفة من الأدلة - تقارير إنجاز المشروعات، وعمليات استعراض برامج الفرص الاستراتيجية القطرية، وتقييمات الأثر، ومؤشرات النتائج الأساسية في نظام متسق يُستخدم لتقييم التقدم المحرز نحو الأهداف الاستراتيجية والتزامات الصندوق في إطار التجديدين الثاني عشر والثالث عشر لموارده. ويتحول تركيز تقييمات الأثر من كونها في المقام الأول أدوات لإعداد التقارير المؤسسية إلى أدوات للتعلم الاستراتيجي والإدارة التكيفية.
- 48- وخلص التقييم الذاتي للإدارة إلى أن وظيفة التقييم الذاتي في الصندوق أصبحت أكثر نضجا الآن مما كانت عليه في وقت استعراض الأقران الخارجي لعام 2019: فقد ساهمت سياسة التقييم المعدلة، وإطار الفعالية الإنمائية 2.0، وتعزيز مكتب الفعالية الإنمائية، والتحسينات في تقارير الإشراف، وتحسين معايير ضمان الجودة لتقارير إنجاز المشروعات وعمليات استعراض إنجاز الاستراتيجية القطرية،¹² والدمج المتزايد لتقييمات الأثر، في تعزيز الاتساق المنهجي وبيان الأدوار والمسؤوليات بشكل أوضح. ويشارك فريق استعراض الأقران الخارجي هذا الاستنتاج.
- 49- ووفقا للتقييم الذاتي، ساهمت أدوات الإبلاغ المؤسسية في الصندوق - ولا سيما التقرير السنوي عن الفعالية الإنمائية واستعراض منتصف المدة لالتزامات تجديد الموارد - في تعزيز الحافز لدى التقييم الذاتي لتوفير بيانات لا تتسم بالدقة فحسب، بل أيضا بالملاءمة الاستراتيجية.
- 50- ويعرض تقرير الفعالية الإنمائية للصندوق أداء الصندوق خلال سنة معينة ويركز على المؤشرات والأهداف من إطار إدارة النتائج. ويبلغ تقرير الفعالية الإنمائية للصندوق بطريقة شاملة عن مجموعة من المؤشرات ابتداء من الموارد البشرية وانتهاء بالكفاءة المؤسسية، وأداء البرامج القطرية، والنتائج على مستوى المشروعات وإدارة الحافظة. وكما هو الحال مع تقارير مماثلة في مؤسسات مالية دولية أخرى، يركز تقرير الفعالية الإنمائية للصندوق على الأداء الإيجابي للصندوق وإنجازاته، مع مناقشة محدودة لمواطن الضعف والتحديات. ومن شأن تقديم عرض أكثر توازنا أن يعزز مصداقية التقرير.
- 51- **وتظهر نتائج المقابلات والاستقصاء شواغل بشأن الجودة المتغيرة لبيانات التقييم الذاتي.** ووفقا للتقييم الذاتي الذي أجراه مكتب الفعالية الإنمائية، فقد تحسنت المواءمة مع المعايير الدولية بشكل كبير. وتنعكس مبادئ فريق الأمم المتحدة المعني بالتقييم بشأن التقييم والشفافية والفائدة في صيغ الإبلاغ الموحدة والتوجيهات المعززة. ويجري تفعيل معايير لجنة المساعدة الإنمائية التابعة لمنظمة التعاون والتنمية في الميدان الاقتصادي في أعمال الإشراف،

¹² عمليات استعراض إنجاز الاستراتيجية القطرية.

وتقارير إنجاز المشروعات، وعمليات استعراض إنجاز الاستراتيجيات القطرية، بينما يُسترشد بمبادئ فريق التعاون في مجال التقييم في الاتساق المنهجي. ومع ذلك، لا يزال يتعين معالجة التفاوت في التنفيذ في الأقاليم - المدفوع بتنوع قدرات الشركاء والشعب، والموارد البشرية، والواقع التشغيلي - من أجل إضفاء الطابع المؤسسي بشكل كامل.

52- وتقر الإدارة بأن فعاليتها المستقبلية ستعتمد على قدرة الصندوق على التركيز على التوحيد والجودة والتعلم والاستخدام. وبدلاً من إدخال مزيد من الإصلاحات، ينبغي التركيز على التطبيق المتسق، وضمان الجودة، والمتابعة الإدارية. ويوافق فريق استعراض الأقران الخارجي بالكامل على هذا النهج.

53- وتشمل الشروط التمكينية الرئيسية استمرارية القيادة من قبل مكتب الفعالية الإنمائية، ووضوح ترتيبات المساءلة مع الشعب التشغيلية، ومواءمة التوجيه المؤسسي والواقع التشغيلي على الأرض. وعلاوة على ذلك، يبرز التقييم الذاتي الحاجة إلى إضفاء الطابع المؤسسي على دورات التعلم التي تترجم أدلة التقييم الذاتي بشكل منهجي إلى عملية صنع القرار. كما يشير إلى فرص توسيع الاستخدام الاستراتيجي للأدوات الرقمية وتحليلات البيانات والتكنولوجيات الناشئة لتعزيز توليد الأدلة وتولييفها واستخدامها. ويدعم فريق استعراض الأقران الخارجي هذه الاعتبارات، بما في ذلك الحاجة إلى الاستفادة من الذكاء الاصطناعي لتحسين الكفاءة والفعالية.

النقاط الأساسية

يُنظر إلى تقييمات مكتب التقييم المستقل على نطاق واسع على أنها ذات مصداقية وسليمة منهجياً، مع تحسّن في الكفاءة. ومع ذلك، تبرز الحاجة إلى تعزيز تحديد الأولويات الاستراتيجية، ومشاركة أصحاب المصلحة، وجودة البيانات، والتوجه نحو التعلم، وإعداد تقارير أكثر تركيزاً وتأثيراً لتحقيق أقصى قدر من المواءمة والاستخدام والأثر المؤسسي.

وأصبحت مصداقية التقييم الذاتي الآن أفضل تنظيماً وأكثر نضجاً مما كانت عليه في عام 2019، مع تطورات واعدة في المستقبل.

3-2 الاستقلالية والحوكمة: وظيفة راسخة ومحترمة، مع الحاجة إلى ضمان الحفاظ على الاستقلالية

54- يتمثل السؤال الرئيسي في إطار معايير الاستقلالية فيما إذا كان مكتب التقييم المستقل يعمل وفقاً لتعريف الاستقلالية (تنظيمياً، وسلوكياً، والحماية من التدخل الخارجي، وتجنب تضارب المصالح). وهذا يترتب عليه آثار على إعداد الميزانية، وتخطيط العمل، وتخصيص الموارد البشرية. ولا يعني ذلك العزلة، بل يعني مستوى موثوقاً من التفاعل والعلاقة مع أصحاب المصلحة.

55- يتمتع مكتب التقييم المستقل بالاستقلالية. ولا تثير نتائج استعراض الأقران الخارجي أي شواغل جوهرية بشأن استقلالية مكتب التقييم المستقل في جميع أبعاد الاستقلالية. بل على العكس، تُعدّ الاستقلالية على نطاق واسع إحدى أبرز نقاط قوة مكتب التقييم المستقل. وقد علّق أحد المجيبين على الاستقصاء قائلاً: "أظهر مكتب التقييم المستقل في ظل قيادته الحالية استقلالية وصرامة كاملتين".

56- وتتضمن سياسة التقييم وصفاً واضحاً للعوامل الهيكلية والسلوكية الأساسية اللازمة لضمان الاستقلالية. ويقدم مكتب التقييم المستقل تقاريره مباشرة إلى المجلس التنفيذي، دون وجود أي خط إبلاغ للإدارة. ولمكتب التقييم

المستقل ميزانية و خطة عمل منفصلتان تستعرضهما لجنة التقييم ويوافق عليهما المجلس التنفيذي. ويتمتع مكتب التقييم المستقل بسلطة اتخاذ القرار النهائي بشأن برنامج عمله ومحتوى جميع منتجات التقييم.

57- ويرى فريق استعراض الأقران الخارجي أن سياسة التقييم المعدلة لعام 2021 تتماشى مع المعايير الدولية والممارسات الشائعة بين المنظمات المماثلة. وتركز السياسة على المبادئ الاستراتيجية الرفيعة المستوى، ولا تقدم، كما ورد في الفقرة 17 من السياسة، تفاصيل المنتجات والعمليات وأساليب التقييم التي قد تحتاج إلى تحديث متكرر. كما أنها لا تقدم تفاصيل عن نوع التقييمات (المستقلة والذاتية) أو المنهجيات والنهج الأساسية التي ينبغي أن تُترك لتقدير مديري الوحدات، وأن تحترم بشكل خاص الخيارات المستقلة والمدرسة لرئيس مكتب التقييم المستقل بعد المشاورات. وعلى النحو المبين في الفصل 1-2، يجد فريق استعراض الأقران الخارجي أن هذه السياسة مرنة بما فيه الكفاية لتشكّل إطاراً لمزيد من التحسينات في وظيفة التقييم. وباستثناء بعض التوضيحات في الملاحق، فإن السياسة نفسها لا تتطلب أي تعديل.

58- ويوصي فريق استعراض الأقران الخارجي بشدة بأن يؤكد الصندوق ليس فقط التزامه بسياسة التقييم لعام 2021 فحسب، بل أن يتخذ أيضاً خطوات ملموسة لحماية الاستقلالية في الممارسة العملية.

النقاط الأساسية

يُنظر إلى مكتب التقييم المستقل على نطاق واسع على أنه مستقل، ولديه ضمانات هيكلية قوية، وإبلاغ مباشر للمجلس، وميزانية منفصلة، وسلطة الكاملة على عمله ومخرجات التقييم الخاصة به. ويُعدّ الحفاظ على استقلالية وظيفة التقييم أمراً أساسياً لمصادقية الصندوق. وتُعتبر سياسة التقييم لعام 2021 مناسبة ولا تتطلب تنقيحاً جوهرياً. وتبرز الحاجة إلى توضيحات موجهة في الملاحق.

4-2 الفائدة: قيمة معترف بها، ولكن الاستفادة منها محدودة

59- تسلط معايير الجدوى الضوء على ملائمة وظيفة التقييم وفائدتها (سواء التقييم المستقل أو التقييم الذاتي) فيما يتعلق بتحسين الأداء المؤسسي للصندوق، لتلبية احتياجات الإدارة من المدخلات الاستراتيجية والتعلم للمستفيدين النهائيين من الصندوق. كما تركز على النشر والتواصل المرتبطين بالتقييم.

60- وتُعتبر الفائدة بالنسبة لمكتب التقييم المستقل جيدة عموماً، مع بعض الاستثناءات. ويعتمد تأثير التقييم على عوامل متعددة، منها العلاقات مع الإدارة فيما يتعلق بتبني التوصيات، وصحتها، وتفصيلها، وحسن توقيتها. وتظهر النتائج أن إدارة وعمليات الصندوق تعتبر عموماً عمل مكتب التقييم المستقل ملائماً ومفيداً.

61- ويتلقى مكتب التقييم المستقل تعقيبات إيجابية متكررة من أصحاب المصلحة الحكوميين بشأن فائدة التقييمات، على شكل مراسلات غير رسمية مكتوبة شفوية وتعقيبات في حلقات عمل أصحاب المصلحة. وقد زاد الاستخدام المباشر لنتائج وتوصيات تقييم الاستراتيجية القطرية والبرنامج القطري في إعداد برامج فرص استراتيجية قطرية جديدة نتيجة للارتباط الوثيق بين تقييم الاستراتيجية القطرية والبرنامج القطري وبرنامج الفرص الاستراتيجية القطرية. وفي عامي 2025/2026، ستكون 69 في المائة من برامج الفرص الاستراتيجية القطرية مسبقة بتقييم للاستراتيجية القطرية والبرنامج القطري - وهو ما يمثل زيادة من 20 في المائة في 2021. وقد علّق أحد المجيبين على الاستقصاء قائلاً: "لقد استخدمت تقييمات مكتب التقييم المستقل في عملي اليومي، إما لاستعراض الحافطة

كجزء من مكتب ميداني للشعبة الإقليمية (المنصب السابق)، أو لتقييم جودة برنامج فرص استراتيجية قطرية قدم للموافقة عليه (الوظيفة الحالية في مكتب الفعالية الإنمائية)."

62- ويتمثل أحد الآراء المثيرة للاهتمام التي أعرب عنها في بعض المقابلات في أن الخضوع لتقييم يجريه مكتب التقييم المستقل يُعتبر أمراً إيجابياً لدى البعض، لأنه قد يساعد على جذب انتباه الإدارة العليا والمجلس التنفيذي. ويثير آخرون ممن أجريت معهم مقابلات شواغل بشأن قدرة العمليات على الاستجابة لكثرة الطلبات الواردة من مكتب التقييم المستقل فيما يتعلق بعمليات التقييم، وبالتالي فهم يحتاجون إلى تحديد أوضح لأولويات منتجات مكتب التقييم المستقل من حيث الأهمية. وبينما قد تكون التقييمات الفردية مفيدة، فإن كثرتها قد تشكل عبئاً على الموظفين مما يقلل من فائدتها. وقد أدلى أحد المجيبين على الاستقصاء بالتعليق التالي: "لقد رأيت أن الطريقة التي يُشرك بها مكتب التقييم المستقل موظفي الصندوق في التقييم متنوعة جداً، اعتماداً على ما إذا كانوا يعملون على تقييم مواضيعي أو تقييم للاستراتيجية القطرية."

63- ويظهر الاستقصاء أن ما يقرب من نصف المجيبين (47 في المائة) قد اطلعوا على منتجات مكتب التقييم المستقل وأفاد 40 في المائة منهم بأنهم استخدموها لإرشاد عملهم. ويستخدم نحو ثلث المجيبين على الاستقصاء منتجات مكتب التقييم المستقل بانتظام (16 في المائة) أو بشكل متكرر جداً (15 في المائة)، بينما يستخدمها نصفهم (49 في المائة) على الأقل عدة مرات. وأفاد 16 في المائة فقط من المشاركين بأنهم لم يستخدموا منتجات مكتب التقييم المستقل مطلقاً خلال الاثني عشر شهراً الماضية.

64- وأعرب أربعة من بين الأعضاء السبعة في المجلس التنفيذي الذين قدموا إجابات كاملة على الاستقصاء عن معرفتهم التامة بعمل وظيفة التقييم في الصندوق، ويمتلك ثلاثة منهم معرفة جزئية بها. ويمثل ستة من هؤلاء الأعضاء بلدان القائمة ألف، أما العضو السابع فهو من بلد من القائمة باء. ولم تكن هناك ردود من بلدان القائمة جيم. وأشار أحد أعضاء المجلس قائلاً: "اطلع على التقييمات بشكل أساسي. ووجدتها جيدة جداً ومتوازنة وواقعية: أعجبتني أنها تهدف إلى تسهيل التعلم وبالتالي تحسين عمل الصندوق بشكل أكبر."

65- إمكانية زيادة التعاون الخارجي والاعتماد المتبادل من أجل تعزيز الفائدة. يُعدّ التعاون مع وظائف التقييم المستقل الأخرى من خلال شبكات التقييم أحد الاحتمالات الممكنة لزيادة جودة التقييمات وملاءمتها وفائدتها. ومكتب التقييم المستقل عضو في كل من فريق التعاون في مجال التقييم، وفريق الأمم المتحدة المعني بالتقييم. كما أنه يعطي حالياً الأولوية لمشاركته في فريق التعاون في مجال التقييم. ويكرس فريق التعاون في مجال التقييم جهوده لتنسيق أعمال التقييم بين المصارف الإنمائية المتعددة الأطراف من خلال العمل على تعزيز استخدام التقييم؛ وتوفير منتدى لتبادل الدروس المستفادة؛ ووضع مؤشرات أداء متناسقة ومنهجيات ونهج للتقييم؛ وتعزيز احترافية التقييم؛ والمساعدة على بناء القدرات في مجال التقييم.

66- ويهدف إطار الاعتماد المتبادل، وهو أحد مبادرات التعاون بين المصارف الإنمائية، إلى تقليل التكرار وزيادة التعاون لتحقيق كفاءة وتنسيق أكثر سلاسة، بما في ذلك فيما يتعلق بالتقييم والتطوير واستخدام أدوات الذكاء الاصطناعي. ويوفر ذلك إمكانيات لتعزيز الفائدة بالنسبة لمكتب التقييم المستقل.

67- وعلى الرغم من اعتبار توصيات مكتب التقييم المستقل مفيدة، فقد أثارت انتقادات بشأنها. ووفقاً للوثائق التوجيهية، يقدم مكتب التقييم المستقل عادةً من أربع إلى ست توصيات. ومع ذلك، غالباً ما تكون التوصيات متعددة الأجزاء وشاملة، وهناك جدل مستمر حول مستوى المرونة مقابل الطابع التوجيهي المطلوب. وتؤكد المقابلات هذا الرأي. فبينما أفاد معظم المجيبين على الاستقصاء بأن التوصيات مفيدة (اعتبرها 39 في المائة مفيدة جداً لعملهم، واعتبرها 43 في المائة مفيدة إلى حد ما)، أبدى البعض أيضاً تعليقات نقدية.

- 68- ويتعلق أحد الانتقادات بعمومية التوصيات. وأدلى المجيبون على الاستقصاء بالتعليقات التالية: "في تقييم حديث أجراه مكتب التقييم المستقل للاستراتيجية القطرية، كانت النتائج والتوصيات عامة جدا ولم تقدم دعما يُذكر لإعداد الاستراتيجية القطرية التالية. وسيكون من المفيد أن تكون التوصيات أكثر تحديدا."
- 69- "تبدو بعض توصيات مكتب التقييم المستقل نظرية وغير عملية في سياقات معينة."
- 70- "أحيانا تكون التوصيات عامة جدا ولا تنطبق فعليا على المشروع الذي يجري تقييمه."
- 71- وهناك انتقاد آخر يتعلق بعدد التوصيات والتوصيات الفرعية. وخلص استعراض جودة تقارير مكتب التقييم المستقل (الذيل الخامس) إلى أنه رغم أن التوصيات قائمة عموما على الأدلة، هناك ميل إلى تجميع ما يمكن وصفه بـ "حزم التوصيات". وعادة ما تتضمن التقييمات من أربع إلى سبع توصيات شاملة، غير أن نحو نصف التقييمات تضمنت أيضا ما يصل إلى 22 توصية فرعية. وبينما يُعدّ تحقيق التوازن بين خصوصية واتساع التوصيات أمرا صعبا بطبيعته، فإن نهج "حزم التوصيات" قد يؤدي إلى إضعاف الرسائل الرئيسية للتقييم وربما يولد حالة من الارتباك بشأن الأولويات والتركيز. وورد إلى استعراض الأقران الخارجي العديد من التعليقات النقدية حول هذا الموضوع. ويفيد إحدى هذه الآراء بأن العديد من التوصيات الفرعية تجعل من الصعب فك تشابكها. وأشار أحد المجيبين على الاستقصاء: "في وقت آخر استعراض أقران خارجي، بدأ مكتب التقييم المستقل في تقليل عدد التوصيات لتسليط الضوء على التوصيات الرئيسية. ولكن في السنوات الأخيرة، زاد مكتب التقييم المستقل عدد التوصيات من خلال التوصيات الفرعية. ويجري تقسيم خمس توصيات رئيسية إلى 25 توصية فرعية، وهو ما يجعل من الصعب تنفيذها. ويبدو أن العديد من التقييمات المواضيعية تقدم توصيات جرى نسخها ولصقها. وغالبا ما تكون هناك (توصية) بزيادة الميزانية والموارد البشرية في مجالات التغذية، والمسائل الجنسانية، والشباب، وما إلى ذلك."
- 72- ويتمثل أحد المقترحات المحددة والقابلة للتنفيذ في أن يعقد مكتب التقييم المستقل اجتماعا تقنيا بينه وبين الإدارة قبل إنجاز التوصيات. وسيوفر ذلك مساحة للتفاهم المتبادل والاتفاق على "ما هو مطلوب"- دون المساس باستقلالية التقييم، وإنما كآلية لضمان أن تكون التوصيات واضحة ومفهومة بشكل مشترك وقابلة للتنفيذ.
- 73- ويتتبع تقرير رئيس الصندوق عن وضع تنفيذ توصيات التقييم وتدابير الإدارة متابعة التوصيات، لكنه يفتقر إلى التحقق الموثوق باستثناء موافقة مكتب التقييم المستقل والتقييم النوعي للتنفيذ.¹³ ويجري تتبع التوصيات سنويا من خلال تقرير رئيس الصندوق عن وضع تنفيذ توصيات التقييم وتدابير الإدارة، الذي تعدّه إدارة الصندوق ويُقدم للجنة التقييم والمجلس التنفيذي. وبما أن مسؤولية تتبع التوصيات تقع على عاتق إدارة الصندوق، فلا توجد آلية للتحقق المستقل أو التقييم النوعي أو المنهجي لنتائج التنفيذ. كما توجد تحديات في تفسير ما ينطوي عليه تقرير التنفيذ.
- 74- ويمكن تعزيز تأثير التقرير السنوي عن التقييم المستقل في الصندوق. فقد تغير التقرير السنوي الرئيسي لمكتب التقييم المستقل خلال فترة استعراض الأقران الخارجي من حيث المحتوى والطول (من 111 صفحة بالإضافة إلى الملحق في 2019 إلى 39 صفحة بالإضافة إلى الملحق في 2025). وأعيد تقديم التقرير السنوي السابق لنتائج وأثر عمليات الصندوق وأعيدت تسميته ليصبح التقرير السنوي عن التقييم المستقل في الصندوق لعام 2022 من أجل تزويد المجلس التنفيذي والإدارة بوصف أكثر شمولاً لأنشطة التقييم التي يضطلع بها مكتب التقييم المستقل ونتائجها والدروس المستفادة منها. ومع استمرار نهج مقارنة النتائج في التقييمات وعرض وتحليل سلاسل زمنية من التصنيفات لتحديد اتجاهات الأداء، يركز النهج الجديد على النتائج الموضوعية. وبحث التقرير السنوي عن

¹³ يمكن تكليف شريك خارجي بهذا الأمر كل سنتين أو ثلاث سنوات.

التقييم المستقل في الصندوق لعام 2024 في موضوعين رئيسيين: العلاقة بين التمويل المشترك وأداء المشروعات وتدخلات التمويل الريفي الشاملة. ويبحث التقرير السنوي عن التقييم المستقل في الصندوق لعام 2025 في استدامة فوائد المشروعات التي يمولها الصندوق. وعلى الرغم من تحسن تأثير واستخدام التقرير من عام 2019 إلى عام 2025، هناك إمكانية لتعزيز تأثير التقرير السنوي عن التقييم المستقل من خلال تحسين النتائج والدروس الرئيسية بشكل أكبر.

75- ويمكن تقديم تقرير أكثر إيجازاً وتركيزاً إلى المجلس التنفيذي ليكون أساساً لدورة سنوية بشأن نتائج التقييم وتوصياته.

76- **تباين في الآراء بشأن فائدة تقارير إنجاز المشروعات.** فمعظم المقيمين على الاستقصاء يرون أن تقارير إنجاز المشروعات مفيدة في تقييم إنجازات المشروعات، وتسهيل تبادل الخبرات والرؤى بين أصحاب المصلحة في المشروع، والمساهمة في تحسين تصميم وتنفيذ المشروعات في المستقبل. ويتضح أحد التعليقات النقدية من خلال الاقتباس التالي: "أعتقد أن تقارير إنجاز المشروعات جيدة جداً، لكن التوقيت أحياناً يكون غير مناسب. ولا أرى جدوى من إعداد تقرير إنجاز مشروع لمشروعات عند اقترابها من تاريخ الإنجاز، وذلك عندما تكون معظم الأنشطة غير مكتملة أو عندما يسعى المشروع إلى التمديد. ويجري تمديد بعض المشروعات لنحو 5 أشهر قبل إغلاقها، ما يتيح لها تنفيذ ما يصل أحياناً إلى 30-40 في المائة من الأنشطة. وهذا لا يسمح للمقيمين بالحصول على معلومات كاملة عن المشروع."

77- **ووفقاً لتقرير الفعالية الإنمائية للصندوق لعام 2025،** قدمت الإدارة آليات أقوى لضمان جودة تقارير إنجاز المشروعات. وبالإضافة إلى ذلك، استحدثت الشعب الإقليمية آليات أكثر صرامة لمراقبة الجودة تتضمن ترتيبات معززة لاستعراض الأقران. وبخلاف الأحكام السابقة، توافق إدارة الصندوق على تقارير إنجاز المشروعات ويمكنها طلب مراجعة التصنيفات في الحالات التي لا يكون فيها المبرر قوياً بما فيه الكفاية.

النقاط الأساسية

تحظى تقييمات مكتب التقييم المستقل بتقدير واسع وتستخدم بشكل متزايد، لكن فائدتها مقيدة بسبب التوصيات التي تتضمن العديد من التوصيات الفرعية، والعبء الملقى على عاتق الموظفين، وضعف التحقق المستقل من المتابعة، وعدم اتساق توقيت تقارير إنجاز المشروعات، والإمكانات غير المستغلة للتعاون الرامي إلى تعزيز الملاءمة والكفاءة.

ويمكن تقديم تقرير سنوي عن التقييم المستقل في الصندوق يكون موجزاً ومركزاً إلى المجلس التنفيذي ليكون أساساً لدورة سنوية بشأن نتائج التقييم وتوصياته.

ويُنظر عموماً إلى التقييمات الذاتية للإدارة على أنها ملائمة ومفيدة للتعلم وتحسين الأداء، لكن فعاليتها تضعف بسبب العبء الملقى على عاتق الموظفين وتفاوت مستوى المشاركة.

5-2 تكامل التقييم المستقل والتقييم الذاتي: تحسينات مع استمرار بعض التحديات

78- من الناحية النظرية، ينبغي أن تكون التقييمات الذاتية للإدارة والتقييمات المستقلة لمكتب التقييم المستقل متكاملة وأن تتجنب التداخل والتكرار. وهذا يستلزم تقسيماً واضحاً للعمل في مجال الرصد والتقييم. كما يعني ذلك أن التقييمات المستقلة المعقدة يمكن أن تعتمد على البيانات التي تُجمع من خلال التقييم الذاتي لضمان مصداقيتها. وتشمل المتطلبات الأساسية لتحقيق التكامل مبادئ توجيهية مشتركة وتفسير معايير التصنيف ومستوى من الثقة

بين جزئي وظيفة التقييم. وينبغي أن يغطي الجزآن معا جانبا كبيرا من احتياجات الصندوق في مجالي المساءلة والتعلم.

79- **تحسين التعاون بين مكتب التقييم المستقل ومكتب الفعالية الإنمائية.** تشير النتائج إلى علاقة أكثر إنتاجية وتواؤما بين الوظيفتين مقارنة بعام 2019. وتمثل أحد التغييرات المهمة منذ استعراض الأقران الخارجي لعام 2019 في زيادة التعاون بين مكتب التقييم المستقل والإدارة، ولا سيما مكتب الفعالية الإنمائية، في تطوير سياسة التقييم ودليل التقييم في الصندوق. وتُعَدُّ السياسة المعدلة السياسة الأولى التي تشمل التقييم الذاتي، مما يضمن رؤية شاملة ونهجا شاملا لوظيفة التقييم في الصندوق، وإتاحة المزيد من التنسيق والتعاون. وقد جرى تعيين مكتب الفعالية الإنمائية كجهة الاتصال التابعة للإدارة لدى مكتب التقييم المستقل، وهو ينسق ردود الإدارة للتقييمات المؤسسية والمواضيعية. ومكتب الفعالية الإنمائية مدعو إلى المشاركة في جميع المشاورات المتعلقة ببرنامج عمل مكتب التقييم المستقل وميزانيته، وتُحال إليه نسخة من جميع المراسلات الرئيسية المتعلقة بالتقييم. ويقدم مكتب التقييم المستقل بدوره تعليقات على المنتجات السنوية التي يعدها مكتب الفعالية الإنمائية مثل تقرير الفعالية الإنمائية للصندوق وتقرير رئيس الصندوق عن وضع تنفيذ توصيات التقييم وتدبير الإدارة.

80- ويؤكد التقييم الذاتي لمكتب الفعالية الإنمائية على الأهمية الاستراتيجية لتعزيز التكامل بين التقييم الذاتي والتقييم المستقل. ويمكن أن تعمل زيادة المواءمة المنهجية، والتبادل التقني الروتيني، واتساق ضمان الجودة على تعزيز المصداقية، والحد من الازدواجية، وتعزيز التعلم. ويعتمد ذلك على استمرار التعاون بين مكتب الفعالية الإنمائية ومكتب التقييم المستقل، والمعايير المشتركة عند الاقتضاء، وترتيبات الحوكمة التي تحسن التعزيز المتبادل مع الحفاظ على الاستقلالية.

81- وبجد فريق استعراض الأقران الخارجي أن التحديد الأوضح للأدوار وإرساء الأسس المفاهيمية المشتركة قد أتاحا فرصة لبناء نظام تكاملي حقيقي. وتتطلب الاستفادة من هذا الأساس المشترك مواءمة عميقة في التوقعات التحليلية، وحوارا قويا بين مختلف الوظائف خلال المراحل الرئيسية من دورات المشروع والدورات القطرية، ونهجاً منهجياً لتوليف رؤى التقييم بما يدعم التعلم التنظيمي. وتظهر النتائج أن وظيفة التقييم تسير في هذا الاتجاه لكنها لا تزال بحاجة إلى مزيد من التقدم.

82- **الاختلافات في تصنيف أداء المشروع.** تصنف الإدارة ومكتب التقييم المستقل¹⁴ نفس معايير التقييم، باستثناء الأثر، الذي لا يقيمه الإدارة دائما (لم تتضمن 10 من أصل 23 من تقارير إنجاز المشروعات في 2024 تقييمات الأثر). ويحدد التقرير السنوي عن التقييم المستقل في الصندوق لعام 2025 فروقا ذات دلالة إحصائية بين تصنيفات مكتب التقييم المستقل وتقارير إنجاز المشروعات في جميع معايير التقييم. وتُعَدُّ تصنيفات مكتب التقييم المستقل أكثر أهمية في المتوسط من تصنيفات الإدارة. وترتبط أكبر التفاوتات بالملاءمة (-0.46)، يليها توسيع النطاق (-0.43) والمساواة بين الجنسين وتمكين المرأة (-0.39). وقد تحسنت هذه التفاوتات بمرور الوقت. ووفقا لتقرير الفعالية الإنمائية للصندوق لعام 2025، فإن التطبيق التدريجي لدليل التقييم المنقح يحقق نتائج في سد هذه الثغرات. وترى الإدارة أن هذه التفاوتات لا تدعو إلى القلق؛ بل على العكس، فهي تثيري بشكل عام النقاش بشأن النتائج وقياس الفعالية الإنمائية. ويتمثل أحد التفسيرات للتفاوت المعبر عنه في المقابلات هو أن الإدارة ومكتب التقييم المستقل يتبعان أدلة تصنيف مختلفة قليلا، وثمة اختلافات في تفسير كل منهما للمعايير.

83- ويرى فريق استعراض الأقران الخارجي أن التفاوتات في التصنيف طبيعية و"صحية" نظرا إلى اختلاف آراء الإدارة ومكتب التقييم المستقل، غير أن الاختلافات في المبادئ التوجيهية تُعَدُّ إشكالية ويمكن معالجتها بسهولة.

¹⁴ يقوم مكتب التقييم المستقل بالتحقق من صحة تقارير إنجاز المشروعات.

وعلاوة على ذلك، ينبغي حل الاختلافات في تفسير هذه المبادئ التوجيهية من خلال الحوار بين مكتب الفعالية الإنمائية ومكتب التقييم المستقل. والتفاوت الحالي في الصندوق ليس أكبر من التفاوتات القائمة في المؤسسات المالية الدولية الأخرى.

النقاط الأساسية

جرى تحسين التكامل والتعاون بين التقييم الذاتي للإدارة والتقييم المستقل خلال الفترة المشمولة بالاستعراض. ولا تزال التفاوتات في التقييم قائمة، وبعضها يُعزى إلى اختلاف المبادئ التوجيهية. وتبرز الحاجة إلى مزيد من المواءمة والحوار والمعايير المشتركة لتعزيز المساءلة والتعلم.

6-2 الإدارة والموارد البشرية: الحاجة إلى مزيج أوسع من المهارات

- 84- تتناول هذه المعايير قدرة وفعالية وكفاءة وظيفة التقييم الذاتي لدى مكتب التقييم المستقل والإدارة، وتحدد ما هو مطلوب لكي تضطلع وظيفة التقييم بدورها المنشود الآن وفي المستقبل المنظور.
- 85- ويمتلك مكتب التقييم المستقل موارد كافية، لكنه يحتاج إلى مزيج أوسع من المهارات لمواكبة تطور الصندوق نحو تقديم المزيد من الدعم للقطاع الخاص والدعم المالي. وتوافرت لمكتب التقييم المستقل عموماً، وفقاً لتقييمه الذاتي، موارد كافية لتنفيذ برنامج عمله. ومع ذلك، فإن زيادة التغطية في عام 2026، وزيادة عدد التقييمات تجاوزت النمو في الميزانية، وهو ما يتطلب زيادة عبء العمل على الموظفين بالإضافة إلى زيادة الكفاءة. وساعدت إضافة موظفين جدد على تلبية جزء من الطلب. وتُكمل القدرات المتاحة بالاستعانة باستشاريين ومحللين ومتدربين.
- 86- ولمواكبة تداعيات الهيكلية المالية الجديدة للصندوق، وإدخال أدوات استثمار جديدة، وزيادة العمليات غير السيادية، وزيادة استخدام الذكاء الاصطناعي في أساليب العمل، يتعين على مكتب التقييم المستقل تعديل قدراته لتلبية هذه الاحتياجات وتطوير خبرات في مجالات مثل:
- العمليات غير السيادية وأدوات التمويل المختلط؛
 - تطوير القطاع الخاص وتمويل سلسلة القيمة؛
 - تكييف منهجيات لقياس الأثر مع سياقات القطاع الخاص.
- 87- وتشير النتائج إلى أن مكتب التقييم المستقل قد اتخذ بعض الخطوات للتحرك في الاتجاه الصحيح، لكنه لم يصل بعد إلى المستوى المطلوب للتعامل بشكل موثوق مع التحديات الجديدة التي يواجهها الصندوق.
- 88- ولكي تضطلع وظيفة التقييم بدور استراتيجي، يجب أن تطور توجهها الاستباقي بشكل أكبر. وقد يتطلب ذلك أيضاً مهارات مختلفة عن النطاق الحالي. ويعد تعزيز التدريب وتطوير القدرات أمراً مهماً لضمان تحفيز موظفي مكتب التقييم المستقل وتقويتهم تقنياً وتحديثهم مهنيًا.
- 89- وتتطلب مصداقية التقييم الذاتي أكثر من مجرد تحديث المبادئ التوجيهية. ورغم تحديث النظام والمبادئ التوجيهية الخاصة بالتقييم الذاتي لكي تتماشى مع الممارسات الشائعة بين المؤسسات المالية الدولية وتعزز الوضوح المنهجي، فإن التعديل في التطبيق لم يحدث بالكامل بعد. وتشير التقييمات الذاتية للإدارة إلى أن النظام يتطلب حالياً الانتقال من الإصلاح الهيكلي إلى تطبيق متنسق على نطاق المنظمة. ويركز ذلك على تحسين موثوقية ودمج الأدلة عبر دورات المشروع والدورات القطرية والمؤسسية، بدلاً من إجراء المزيد من إعادة التصميم الإجرائي.

- 90- ويُعدّ تفاوت مستويات قدرات الرصد والتقييم بين وحدات إدارة المشروعات والمكاتب القطرية عائقاً. وترى الإدارة أن هذا التفاوت لا يعكس نقصاً في الالتزام، بل ينشأ عن قيود هيكلية - مثل المطالب التشغيلية المتنافسة، أو دوران الموظفين، أو الحاجة إلى دعم عملي أكثر استدامة في تطبيق التوجيهات المنهجية. وتعمل الأفرقة في ظل ظروف موارد ونماذج توظيف وبيئات خارجية مختلفة، مما يؤثر بطبيعة الحال على انتظام وجودة جمع البيانات. وتجد الإدارة في تقييمها الذاتي أدلة على وجود ثغرات مستمرة في توازن عبء العمل، ووقت الموظفين، والمهارات التقنية، والحوافز المتعلقة بالتفكير التقييمي.
- 91- وقد سلط العديد من الموظفين الضوء على أهمية مواصلة الاستثمار في التدريب العملي، والمرافقة المستهدفة، والتوجيه الأكثر وضوحاً الذي يترجم المبادئ التقييمية إلى ممارسات قابلة للتنفيذ ومناسبة للسياق.
- 92- وتشير عدة مقابلات إلى ضعف قدرات الرصد والتقييم في وحدات إدارة المشروعات والحكومات الشريكة باعتباره عائقاً رئيسياً أمام الحصول على جودة كافية لبيانات المشروع. ومن الواضح أنه يتعين تعزيز الجهود الرامية إلى معالجة هذه المسألة.
- 93- ومن الضروري معالجة هذه القيود ليس لتحسين جودة منتجات التقييم الذاتي فحسب، بل أيضاً لدمج المنطق التقييمي في صنع القرارات التشغيلية. ويكتسي تعزيز القدرات على مستوى المكاتب القطرية، حيث غالباً ما تكون تحديات قابلية التقييم أكثر حدة، أهمية خاصة. وتؤكد التقييمات التي تلقاها فريق استعراض الأقران الخارجي من ممثلي البلدان على الحاجة إلى تدريب منهجي في المهارات الأساسية المتعلقة بالتوريد والإدارة المالية والرصد والتقييم ومسائل الموارد البشرية.
- 94- ووفقاً لتقرير الفعالية الإنمائية للصندوق لعام 2025، تولى فريق المعرفة والتعلم في الصندوق (جزء من مكتب الفعالية الإنمائية) في أوائل عام 2025 إدارة أكاديمية العمليات في الصندوق، وهي آلية مؤسسية رئيسية مصممة لتعزيز القدرات التشغيلية للأفرقة القطرية والخبراء التقنيين والشركاء. وتتكون أكاديمية العمليات من ثلاث أدوات رئيسية: دورات التعلم الإلكتروني (المستوى الأول)، والحلقات الدراسية الشبكية التفاعلية وفعاليات التدريب الحضوري (المستوى الثاني)، وبرنامج إرشاد منظم. ويوصى برصد وتقييم الجهود عن كثب بعد فترة من سنتين إلى ثلاث سنوات للتحقق من مدى نجاحها.

النقاط الأساسية

يمتلك مكتب التقييم المستقل الموارد، لكنه يحتاج إلى مهارات أوسع وقدرة استباقية لتلبية متطلبات الصندوق المتطورة، لا سيما في العمليات غير السيادية.

ويمثل اختلاف مستويات قدرات الرصد والتقييم في الأقاليم والمكاتب القطرية مصدر قلق. ويتطلب تعزيز التقييم الذاتي وتطبيقه بشكل متسق توفير التدريب العملي، وبناء قدرات على المستوى القطري، وتحسين دمج الأدلة في صنع القرار.

7-2 بيئة تمكينية للتعليم والتطوير: إمكانات كبيرة لمزيد من التحسين

- 95- تتضمن هذه المعايير تقييماً للثقافة التنظيمية للصندوق، والعوامل التي تدعم وتقيد الاستخدام والتعلم من التقييمات.
- 96- ويُعدّ الصندوق منظمة تعليمية نوعاً ما. وفي التقييم الذاتي لمكتب التقييم المستقل، يمنح موظفوه ثقافة التعلم والتقييم في الصندوق تصنيفاً عالياً. وعلاوة على ذلك، يرى مكتب التقييم المستقل أن ارتفاع الطلب على التقييم من إدارة الصندوق يعكس بيئة تنظيمية إيجابية للتقييم، بالإضافة إلى فهم القيمة التي يضيفها التقييم المستقل على

الاستراتيجيات والبرامج المؤسسية والقطرية. ويدعم الاستقصاء هذه الآراء. وجاء أحد التعليقات على النحو التالي: "تساعد التقييمات الموظفين على الخروج من الروتين اليومي للتفكير في الحاجة وكفاءة ما تقوم به المنظمة حالياً. وينبغي أن تحفز الدروس المستفادة الموظفين والإدارة العليا على استكشاف سبل التحسين."

97- وأفاد معظم المجيبين على الاستقصاء (90 في المائة) أن الصندوق منظمة تعليمية. وتُعدّ هذه النتيجة متسقة خلال فترة الخدمة، لكنها تميل إلى الارتفاع قليلاً بالنسبة للمشاركين الذين لديهم أكثر من ست سنوات من الخبرة العملية في الصندوق. ومع ذلك، فإن الموظفين الإداريين يميلون بدرجة أقل قليلاً إلى اعتبار الصندوق منظمة تعليمية، حيث أفاد 15 في المائة بأن هذا الأمر غير صحيح على الإطلاق.

98- وتقدم نتائج المقابلات صورة أكثر دقة. ولا تستخدم الأدلة دائماً عند التخطيط لمشروعات جديدة. فعلى سبيل المثال، عندما تظهر الأدلة المستقاة من التقييمات أن نوعاً من العمليات لا يحقق النتائج المرجوة، لا يُراعى ذلك دائماً عند التخطيط لعمليات جديدة.

99- ويؤثر حسن توقيت نشر تقارير مكتب التقييم المستقل على التعلم الأوسع نطاقاً. وتظهر النتائج أن التأخيرات في نشر تقارير مكتب التقييم المستقل تمنع الاستخدام الواسع النطاق لنتائج هذه التقارير وتوصياتها. وقد علّق أحد المجيبين على الاستقصاء قائلاً: "غالباً ما تُنشر منتجات مكتب التقييم المستقل مع تأخيرات كبيرة. ويؤثر ذلك على توقيتها ويقلل من تأثيرها المحتمل، حيث قد لا يرى المسودات سوى مجموعة صغيرة من الموظفين الداخليين، بينما يتسنى لمجتمع أصحاب المصلحة الأوسع نطاقاً - بما في ذلك الحكومات وأفرقة المشروعات والشركاء الإنمائيون - الوصول إلى هذه المسودات في وقت لاحق بكثير، وذلك عندما تكون فرص التعلم وتصحيح المسار أكثر محدودة."

100- العوامل التمكينية والمقيدة للتعلم. والعوامل التمكينية للتعلم الأكثر ذكراً من قبل المجيبين هي التعاون والعمل الجماعي في المشروعات والإدارات (71 في المائة) والوصول إلى التدريب والموارد وبرامج التطوير (58 في المائة). وبفيد 41 في المائة من المجيبين بوجود آليات للتعقيبات خاصة بالتعلم.

101- كما ذكر المجيبون مجموعة متنوعة من العوامل التي قد تقيد التعلم في الصندوق، مثل الأولويات السياسية وثقافة المنظمة. وعند مناقشة إدارة المعرفة، ذكر أحدهم: "لم يعد هناك هيكل واضح لإدارة المعرفة في الوقت الحالي، لذلك، لا يشعر أحد بالمسؤولية عن قيادة إدارة المعرفة. فحين أصبحت إدارة المعرفة مسؤولية الجميع، لم تعد في الواقع مسؤولية أحد. كما أن العلاقات بين المقر الرئيسي والميدان في مسألة إدارة المعرفة غير واضحة."

102- وتشير التقييمات الذاتية للإدارة إلى أن الموظفين يسلطون الضوء على أن الاستخدام العملي للأدوات والتوجيهات يعتمد بشكل كبير على مشاركة القيادة، وتوافر الموارد، ومدى إدراك فائدة الأدلة التقييمية في صنع القرارات التشغيلية.

103- ويشير التقييم الذاتي إلى أن الصندوق يتجه نحو ثقافة تقييمية أكثر تماسكاً. ويورد أمثلة مشجعة مثل العيادات المواضيعية وجلسات الإحاطة المنظمة وصيغ تبادل الأقران، لكنه يرى أن هذه الممارسات لا تزال متفاوتة ولم تُرسخ بعد في دورة منهجية تجمع الأدلة بشكل روتيني عبر الإشراف، وتقارير الإنجاز، والاستقصاءات، وتقييمات الأثر. وتستمر الفوارق بين الأقاليم وأنواع المشروعات ومستويات القدرات في مجال الرصد والتقييم. وعلى الرغم من تحسن نظم الإبلاغ، لا تزال عمليات التعلم متفاوتة وغالباً ما تعتمد على المبادرة الفردية بدلاً من الممارسات المؤسسية.

- 104- ويتفق فريق استعراض الأقران الخارجي مع الإدارة على أن تعزيز البيئة التمكينية - من خلال بناء القدرات الموجهة، وتسلسل إنتاج الأدلة، ودورات التعلم المؤسسية المقصودة - أمر بالغ الأهمية لضمان استمرار الاسترشاد بالأدلة التقييمية في تصميم المشروعات والإشراف، والتموضع الاستراتيجي.
- 105- وترى الإدارة أيضا أن مواعمة توقيت منتجات التقييم الذاتي مع نقاط القرارات التشغيلية الرئيسية - لا سيما تصميم المشروعات، ودورات الإشراف، وعمليات استعراض برنامج الفرص الاستراتيجية القطرية - من شأنها أن تضمن توافر الأدلة التقييمية عند اتخاذ القرارات.
- 106- وسيساعد استخدام الذكاء الاصطناعي لمعالجة كميات كبيرة من البيانات التقييمية - مثل تقارير إنجاز المشروعات، وسجلات الإشراف، ومجموعات بيانات الرصد والتقييم - على تمكين الصندوق من تجميع الدروس المستفادة بكفاءة أكبر، وتصنيفها حسب المجال المواضيعي أو نوع المشروع أو الجغرافيا، وتحويل المعلومات الموزعة إلى معرفة قابلة للتنفيذ.
- 107- وتشمل فوائد تعزيز المواعمة بين التقييم الذاتي للإدارة والتقييم المستقل لمكتب التقييم المستقل زيادة مصداقية الأدلة التقييمية، وبيئة تقييمية أكثر قابلية للتنبؤ، وثقافة أقوى للمساءلة والتعلم.
- 108- **ويطبق مكتب التقييم المستقل نهجا قائما على علم الأعصاب في التواصل.** فقد استخدم مكتب التقييم المستقل نهجا قائما على علم الأعصاب في تعامله مع أصحاب المصلحة في الصندوق، مدركا أن أنماط التواصل قد يكون لها أثر كبير على مدى تقبل النتائج، لا سيما عندما تكون نقدية أو سلبية. ويركز هذا النهج على استخدام التعاطف والرحمة لتهيئة مساحات آمنة نفسيا للتعبير ومشاركة وجهات نظر غير متحيزة، وإدارة التعقيد، وفهم الدوافع الجوهرية، واستخدام أطر تركز على المستقبل.
- 109- وتشير تجربة أعضاء فريق استعراض الأقران الخارجي إلى أنه يفضل اتباع نهج شامل للتواصل من خلال النظر إلى المحتوى والعلاقات والسياق والقنوات معا، بما يضمن عمل جميع العناصر بصورة متسقة. ويكتسي حسن توقيت النشر والتواصل أهمية خاصة في بيئة تنظيمية معقدة مثل الصندوق.
- 110- **تنوع استثنائي في التواصل لدى مكتب التقييم المستقل.** أعجب فريق استعراض الأقران الخارجي بالتنوع الكبير في منتجات النشر والتواصل التي يقدمها مكتب التقييم المستقل، وهو ما يتجاوز ما تقوم به معظم إدارات التقييم في المؤسسات المالية الدولية الأخرى. ويرى أعضاء المجلس الذين أجابوا على الاستقصاء أن مكتب التقييم المستقل فعال أو فعال للغاية في نشر نتائج التقييم والتواصل بشأنها، وكذلك في تسهيل تبادل الخبرات بين أصحاب المصلحة في المشروع وتعزيز تعلم الدروس. وأعضاء المجلس أقل دراية بالتواصل الذي يقوم به مكتب التقييم المستقل مع البلدان الشريكة، حيث لم يتمكن ثلاثة أعضاء من تقييم هذه الجوانب.
- 111- وتجاوز عدد زيارات الموقع الشبكي لمكتب التقييم المستقل 400 000 زيارة. وقد أجري تمرين باستخدام الذكاء الاصطناعي لتحليل اتجاهات التفاعل مع مكتب التقييم المستقل على وسائل التواصل الاجتماعي خلال الفترة من يناير/كانون الثاني إلى يوليو/تموز 2025. وحلّل التمرين التفاعل على منصتي التواصل الاجتماعي الخاصتين بمكتب التقييم المستقل (منصة X وLinkedIn) اللتين تضمّان نحو 3 600 و8 000 متابع على التوالي.
- 112- **تباين في الآراء بشأن القدرة الاستيعابية المتعلقة بالتعلم في الصندوق.** القدرة الاستيعابية هي قدرة المنظمة على تحديد واستيعاب وتحويل واستخدام النتائج والدروس المستفادة من وظيفة التقييم وغيرها من مصادر الأدلة. وتعتمد القدرة الاستيعابية على الانفتاح على المعرفة الجديدة، وكذلك على القدرة على الاستجابة وحسن توقيت الاطلاع على النتائج والدروس المستفادة والحافز للقيام بذلك. وبالنسبة لوظيفة التقييم، من الصعب إيجاد التوازن الصحيح بين القليل جدا والكثير جدا عندما يتعلق الأمر بالتقييمات وعمليات التقييم. ويرى موظفو مكتب التقييم المستقل أن

القدرة الاستيعابية للصندوق كافية، ويذكرون أن إجراء تقييم أو تقييمين من التقييمات المؤسسية أو المواضيعية سنويا ينبغي أن يكون ضمن قدرة الصندوق على الاستجابة والاستيعاب. ويشكك العديد ممن أجريت معهم مقابلات في انطباق هذا الأمر، ويجدون أن القدرة الاستيعابية للإدارة تشكل تحدياً. والحجة هي أن مكتب التقييم المستقل يتمتع بموارد أفضل مقارنة بأجزاء أخرى من المنظمة. ويؤدي العدد الكبير من التقييمات إلى تقويض القدرة على الاستجابة للطلبات أثناء عمليات التقييم والتعامل مع الدروس والتوصيات المستمدة من التقييمات.

113- **الاستخدام المسؤول والاستراتيجي للذكاء الاصطناعي في التقييم المستقل.** من المرجح أن يؤدي الذكاء الاصطناعي دوراً متزايد الأهمية في مستقبل التقييم المستقل، لا سيما في تحليل الأدلة النوعية والكمية الكبيرة والمعقدة، وأيضاً لتحسين الوصول إلى رؤية التقييم من قبل جميع أعضاء المنظمة لتغذية تصميم المشروعات بطريقة منهجية و"إلزامية".

114- وبينما يمكن لهذه الأدوات دعم المقيمين بشكل كبير، فإن استخدامها يثير أيضاً أسئلة مهمة تتعلق بالتحيز، والشفافية، وقابلية التفسير، وحماية البيانات، والاعتماد المفرط على المخرجات المؤتمتة. ومن الضروري معالجة هذه المخاطر لحماية استقلالية التقييم ومصداقيته وموثوقيته. ولهذا الغرض، هناك جدوى في وضع توجيهات واضحة بشأن الاستخدام الأخلاقي والمناسب للذكاء الاصطناعي في التقييم، مع الاعتماد على العمل الجاري مع المؤسسات النظرية، بما في ذلك المؤسسات المالية الدولية الأخرى وفريق التعاون في مجال التقييم وفريق الأمم المتحدة المعني بالتقييم والشبكات ذات الصلة، والمساهمة في هذا العمل. ومن شأن هذا التنسيق أن يساعد على تعزيز الاتساق مع الممارسات الدولية الجيدة الناشئة مع تجنب التهجس المجزأة أو غير المتسقة.

115- ويجد فريق استعراض الأقران الخارجي أن "استراتيجية دمج الذكاء الاصطناعي في مكتب التقييم المستقل للفترة 2024-2026" هي بداية جيدة لتوفير إطار منظم لترتيب حالات الاستخدام بحسب الأولوية، وإدارة المخاطر، وتوضيح ترتيبات الحوكمة، وضمان المواءمة مع معايير وقيم التقييم. وينبغي تشجيع استخدام الذكاء الاصطناعي من خلال توفير استثمارات تكميلية في تدريب الموظفين والإلمام بالذكاء الاصطناعي. وهذا الأمر ضروري لضمان بقاء المقيمين على اطلاع ومستخدمين ناقدين لهذه الأدوات، بحيث يعمل الذكاء الاصطناعي على تعزيز - بدلاً من استبدال - الحكم المهني والإشراف البشري.

116- ويمثل الذكاء الاصطناعي فرصة استراتيجية لوظيفة التقييم - لتحسين التغطية، والتوقيت، وسهولة الوصول إلى الأدلة، والكفاءة. ولكنه يتطلب حوكمة مقصودة تشمل وضع مبادئ توجيهية أخلاقية واضحة، وتطوير قدرات الموظفين، والتنسيق مع المؤسسات النظرية. وينبغي أن يعزز الذكاء الاصطناعي الحكم التقييمي المهني، لا أن يستبدله. وتُعد استراتيجية مكتب التقييم المستقل الحالية أساساً قوياً؛ كما أن التنفيذ وتعزيز الإلمام لدى الموظفين من الخطوات ذات الأولوية التالية.

117- **الأدوار، والمسؤوليات، وضرورة الاستقلالية في التقييم.** وتضع سياسة التقييم في الصندوق لعام 2021 إطاراً قوياً للتقييم المستقل، حيث تحدد بوضوح أدوار ومسؤوليات المجلس التنفيذي، ولجنة التقييم، وإدارة الصندوق، ومكتب التقييم المستقل. ويكلف المجلس التنفيذي ولجنة التقييم برصد تنفيذ هذه السياسة، لضمان توافر الموارد الكافية للتقييم المستقل، والاعتراف به، ومتابعته بشكل منهجي. كما يتحمل المجلس المسؤولية الحاسمة عن تعيين مدير مكتب التقييم المستقل من خلال عملية مفصلة وشفافة، كما هو موضح في ملحق السياسة. وإدارة الصندوق مسؤولة بدورها عن حماية استقلالية مكتب التقييم المستقل، والمشاركة في التشاور المنتظم معه، وتوفير الوصول الكامل وغير المقيد إلى المعلومات والدعم اللازم للتقييمات. ومكتب التقييم المستقل، الذي يُقدم تقاريره مباشرة إلى المجلس التنفيذي، مسؤول عن الالتزام بالأعراف والمعايير الدولية التي يحددها فريق التعاون في مجال التقييم

وفريق الأمم المتحدة المعني بالتقييم وشبكة التقييم التابعة للجنة المساعدة الانمائية التابعة لمنظمة التعاون والتنمية في الميدان الاقتصادي.

118- ومع ذلك، فإن السياسة المصممة جيداً لا تكفي وحدها لضمان استقلالية مكتب التقييم المستقل. وتعتمد المصادقية والفعالية على الحفاظ العملي على الاستقلالية في جميع العمليات، بما في ذلك تطوير برامج العمل، وتنفيذ ونشر التقييم، وإدارة الميزانية والموارد البشرية وتعيين الموظفين. ويجب الحفاظ على الاستقلالية بنشاط في العمليات اليومية، وصنع القرار، والثقافة المؤسسية.

119- وتُعدّ عملية تعيين مدير مكتب التقييم المستقل حجر الزاوية لاستقلالية المكتب. ويجب أن تكون هذه العملية شفافة بالكامل، وقائمة على الجدارة، وخالية من أي تأثيرات خارجية. ولتجنب تضارب المصالح وضمان الموضوعية، يجب أن يكون توظيف وإدارة موظفي مكتب التقييم المستقل مستقلين تماماً عن إدارة الصندوق ضمن إطار الموارد البشرية في الصندوق. وبعبارة أخرى، يجب الحفاظ على الاستقلالية التشغيلية لمكتب التقييم المستقل في جميع جوانب عمله، بما في ذلك الوصول غير المقيد إلى المعلومات وحرية النشر والتوزيع.

120- واستقلالية مكتب التقييم المستقل ليست مجرد مطلب إجرائي؛ بل هي أساس الثقة في وظيفة التقييم المستقل في الصندوق. وعندما تتعرض الاستقلالية للخطر - سواء من خلال التدخل في عمليات الموارد البشرية، أو التوظيف المسيس أو القيود على الاستقلالية التشغيلية - تتلاشى مصداقية التقييمات. وهذا بدوره يضعف قدرة الصندوق على التعلم والتكيف وإظهار المساءلة أمام أصحاب المصلحة لديه والمجموعات التي يخدمها.

121- ويختلف التقييم الذاتي الذي تقوده الإدارة بطبيعته عن مكتب التقييم المستقل عندما يتعلق الأمر بالاستقلالية الهيكلية. فعلى الرغم من أن التقييم الذاتي يجريه المسؤولون عن المشروع أو العملية التي سيجري تقييمها، ينبغي أن تكون هذه التقييمات سليمة وخالية من التحيزات. وكما تنص السياسة، فإن المبادئ الستة للتقييم - الفائدة، والحياد والمصادقية، والشفافية، والشراسة، والتشاور والتعاون، والقابلية للتقييم والقيمة مقابل المال/الفعالية من حيث التكلفة - تنطبق على التقييم الذاتي والتقييم المستقل على حد سواء. ويجب أن يكون جزءاً من وظيفة التقييم مكملين لبعضهما البعض لضمان أدلة موثوقة ومناسبة التوقيت للتعلم والمساءلة ومواصلة تحسين عمليات الصندوق.

النقاط الأساسية

يقدر الصندوق التقييم والتعلم، ولكن الأدلة على استخدامهما متفاوتة. وتحد التأخيرات، وفجوات القدرات، وضعف إدارة المعرفة من الاستفادة منهما.

ويمكن أن يساهم تحسين التوقيت والقيادة ودورات التعلم المنهجية والتوليف المدعوم بالذكاء الاصطناعي في تعزيز الأثر.

ونظراً إلى تزايد استخدام الذكاء الاصطناعي، يحتاج مكتب التقييم المستقل إلى استراتيجية واضحة وتوجيهات صريحة لضمان استخدام الذكاء الاصطناعي بشكل متنسق ومسؤول وبما يتوافق مع مبادئ التقييم والاستقلالية والمساءلة.

3- التوصيات

122- يتمثل الهدف العام من التوصيات في زيادة المواءمة بين وظيفة التقييم واستراتيجية الصندوق وعمليات الإصلاح فيه. ولدى وظيفة التقييم القدرة على توفير المزيد من الأدلة والمعرفة لعمليات ملائمة استراتيجيا مثل مبادرات تجديد الموارد والإصلاح. والمواءمة تعني التفاعل الوثيق بين مكتب التقييم المستقل ومكتب الفعالية الإنمائية/الإدارة والمجلس التنفيذي. وبالنسبة لوظيفة التقييم، يترتب على ذلك آثار تتعلق بتخطيط العمل، واختيار منتجات التقييم، وحسن توقيت النشر والتواصل. أما بالنسبة للإدارة، فيترتب على ذلك آثار على احتياجات تبادل المعرفة، والاستعداد للبحث عن القرارات وإسنادها إلى الأدلة، والنظر فيها، وتحسينها والتكيف معها. والمواءمة الوثيقة قادرة على تعزيز الفعالية الإنمائية، وفي النهاية تحقيق الأهداف التشغيلية للصندوق. ويمكن أن يؤدي الاستخدام المتزايد والوعي للذكاء الاصطناعي والتكنولوجيات الأخرى دورا محوريا.

123- **التوصية 1:** ينبغي أن يعمل مكتب التقييم المستقل على إعادة موازنة مزيج منتجاته من خلال تحسين الانتقائية، بحيث تركز الموارد على تقييمات مختارة استراتيجيا تعزز فائدة التقييم لصنع القرار في الوقت المناسب وبشكل مستنير.

124- وهذا يعني إعطاء أولوية أوضح للتقييمات الأكثر أهمية استراتيجيا مع مراعاة القدرة الاستيعابية للإدارة للمشاركة في عمليات التقييم وتنفيذ التوصيات. وبينما ينبغي أن يستمع مكتب التقييم المستقل في برنامج عمله لمطالب المجلس والإدارة، ينبغي أن يكون واعيا بنفس القدر للاحتياجات الحالية والمستقبلية من الأدلة التقييمية، بما في ذلك مراعاة الجدول الزمني للتجديد الرابع عشر لموارد الصندوق. وبدلا من الاستمرار في إجراء تقييمات المشروعات وتقييمات الاستراتيجيات القطرية والبرامج القطرية على حوافز المشروعات المحدودة أو عندما لا تكون متبوعة ببرنامج فرص استراتيجية قطرية، ينبغي أن يجري مكتب التقييم المستقل تقييمات جامعة ودراسات توليفية أو أن يُعد مذكرات معلومات بشأن المسائل الشاملة والمناسبة التوقيت، مستفيدا من نتائج التقييم والدروس القائمة لإرشاد أولويات الصندوق وتوجهاته الاستراتيجية، بما في ذلك العمليات غير السيادية. ويترتب على ذلك آثار تتعلق بتخطيط العمل، ومزيج مهارات الموظفين، ومنهجية إشراك أصحاب المصلحة على المستوى القطري، واستخدام التكنولوجيات.

125- **التوصية 2:** ينبغي أن تعمل الإدارة على زيادة الجهود الرامية إلى تحسين التقييم الذاتي والعمل مع مكتب التقييم المستقل على تحسين التكامل مع التقييمات المستقلة، بما يتيح حلقات تعلم أقوى واستخداما أكثر اتساقا لرؤى التقييم في صنع القرار.

126- ومع نضوج التقييم الذاتي من خلال تنفيذ التطوير المخطط له حاليا من قبل مكتب الفعالية الإنمائية كما ورد في التقييم الذاتي لمكتب التقييم المستقل، ينبغي استعراض ومراجعة تقسيم العمل بين مكتب الفعالية الإنمائية ومكتب التقييم المستقل لتجنب التداخل وتعزيز الميزة النسبية لجزأي وظيفة التقييم. كما ينبغي ضمان مزيد من الاتساق في المنهجيات والمعايير من خلال اتخاذ التدابير المناسبة. وينطوي على ذلك آثار تتعلق بالتقييم المستقل والتقييم الذاتي للمشروعات، والبرامج القطرية. وينبغي إجراء تقييم متابعة لوظيفة التقييم الذاتي بعد عامين عندما يكون الإصلاح أكثر تقدما لتقييم مدى ملاءمة وفعالية واتساق النظام الجديد والممارسات والجودة ونتائج وظيفة التقييم الذاتي المنقحة بشكل عام.

127- **التوصية 3:** ينبغي أن يعمل الصندوق على تعزيز ثقافة التعلم لديه من خلال تعزيز التعاون الوثيق بين الإدارة ووظيفة التقييم، وذلك بهدف تحسين الاستفادة من دروس التقييم واستخدامها في القرارات الاستراتيجية والتشغيلية.

128- ويشمل ذلك تحسين حلقات التعلم عبر الصندوق، وإدخال مجالات للتفكير المنهجي، ومراجعة أساليب الإبلاغ عن تنفيذ توصيات مكتب التقييم المستقل والتحقق من ذلك. كما يعني ذلك ضمان تحديد دروس التقييم الرئيسية المستفادة من التقييم واستخدامها في المراحل المبكرة من المشروع وتصميم برنامج الفرص الاستراتيجية القطرية. وينطوي ذلك على الاستخدام الاستباقي لأساليب التعلم في جميع أجزاء الصندوق، ومشاركة أعمق للمجلس التنفيذي، بما في ذلك عقد جلسات مخصصة لمناقشة نسخة محسنة من التقرير السنوي الرئيسي لمكتب التقييم المستقل، أي التقرير السنوي عن التقييم المستقل في الصندوق، وغيره من منتجات التقييم الرئيسية. ويعتمد التعلم المعزز بدرجة أكبر على السلوك والحوافز والعمليات المنظمة جيداً، مقارنة باعتماده على التدابير التعليمية المتطورة والمكلفة.

129- **التوصية 4:** ينبغي أن يعمل مكتب التقييم المستقل على زيادة تركيز ووضوح توصياته لتعزيز قابليتها للتنفيذ وقيمتها في صنع القرار.

130- وسيضمن تعزيز التفاعل الأوثق مع الإدارة في صياغة التوصيات وترتيبها بحسب الأولوية أن تكون التوصيات مفهومة للجميع، وواضحة، ومفيدة، وواقعية، وقابلة للتنفيذ. ولصون مبدأ أن تكون التوصيات مدعومة جيداً بالنتائج، ينبغي لمكتب التقييم المستقل تعزيز الشفافية فيما يتعلق بجودة البيانات والتأليث، وإيلاء اهتمام أكثر منهجية للبيئة، وتعميق التفكير في الصلاحية الخارجية للنتائج. وينبغي تجنب التوصيات الفرعية، كما ينبغي التحقق من رصد التوصيات من خلال إجراء تقييمات مستقلة. وحسن التوقيت عامل ضروري لتعزيز إمكانية الاستفادة من الدروس والتوصيات.

131- ويمكن عقد اجتماع تقني بشأن التوصيات بين مكتب التقييم المستقل والإدارة قبل إنجاز التوصيات، ليس لمناقشة جوهر التوصيات المبنية على الاستقلالية التقييمية، بل كآلية لضمان أن تكون التوصيات مفهومة من الطرفين وواقعية وقابلة للتنفيذ.

132- ويمكن النظر في آلية تحقق مستقلة ومنظمة بشأن تقرير رئيس الصندوق عن وضع تنفيذ توصيات التقييم وتدابير الإدارة من قبل شريك خارجي لتقييم تبني التوصيات ونتائجها.

133- **التوصية 5:** ينبغي أن يعمل مكتب الفعالية الإنمائية ومكتب التقييم المستقل على تعزيز التعاون مع المبادرات العالمية ذات الصلة¹⁵ بشأن تنمية قدرات الرصد والتقييم لتوضيح الأدوار، وتعزيز القدرات، وتحسين قيمة صنع القرار في الرصد والتقييم داخلياً ومع الشركاء القطريين على حد سواء.

134- وتعدّ قدرات الرصد والتقييم فيما بين الموظفين التشغيليين في الصندوق على جميع المستويات وممثلي البلدان الشريكة ووحدات إدارة المشروعات أمراً أساسياً لمصادقية البيانات والقدرة على تنفيذ المشروعات. وينبغي أن يواصل مكتب التقييم المستقل ومكتب الفعالية الإنمائية تعزيز جهودهما في بناء قدرات الرصد والتقييم وبناء تعاون أقوى مع المبادرات العالمية ذات الصلة.

135- **التوصية 6:** ينبغي أن يعيد المجلس التنفيذي تأكيد التزامه بسياسة التقييم لعام 2021 وأن يحمي استقلالية مكتب التقييم المستقل ودور مكتب الفعالية الإنمائية من خلال توضيح الملاحق في السياسة.

136- ويرى استعراض الأقران الخارجي أن السياسة نفسها متوافقة مع أفضل الممارسات في المنظمات المماثلة ومرنة بما يكفي لتوفر إطاراً لمزيد من التحسينات في وظيفة التقييم. وينبغي أن تبين الملاحق أدوار ومسؤوليات مكتب

¹⁵ مثل مبادرة التقييم العالمية، بما في ذلك مركز التعلم بشأن التقييم والنتائج.

الفعالية الإنمائية، وأن توضح أن المجلس مسؤول بالكامل عن اختيار وإقالة مدير التقييم دون تدخل، وأن يخصص وقت كاف في لجنة التقييم لكل تقييم يجري تقديمه.¹⁶

¹⁶ ينبغي أن توضح الملاحق ما يلي: (1) دور ومسؤوليات مكتب الفعالية الإنمائية، (2) تخصيص وقت كاف لكل منتج تقييم يعرض على لجنة التقييم، (3) أن المجلس مسؤول حصرياً عن اختيار وإقالة مدير مكتب التقييم المستقل، (4) أنه ينبغي أن تكون عملية التوظيف قائمة على الجدارة، وشفافة تماماً، وخالية من أي تدخلات خارجية، بما في ذلك إدارة الصندوق، والرئيس، ومدير مكتب التقييم المستقل المنتهية ولايته.

4 Appendices

Appendix 1: EPR Evaluation Framework ([LINK](#))

Appendix 2: EPR approach and methodology

Appendix 3: List of interviewees

Appendix 4: Surveys of staff and members of the Executive Board

Appendix 5: Assessment of IOE evaluation reports

Appendix 1: EPR Evaluation Framework ([LINK](#))

Appendix 2: EPR approach and methodology

The Evaluation Framework (submitted by the EPR in November 25 and approved by IOE and Management) was developed in response to the approach paper (9 October 2025) and recognizes guidelines for such reviews (ECG, UNEG and OECD/DAC) including conversations with selected key stakeholders and an initial identification of issues related to the evaluation function in the current and emerging internal and external context. It was also based on the panel members' broad expertise and experience and the time and resources available to complete the EPR. The findings and recommendations of the 2019 EPR formed the basis for this EPR.

The EPR has also assessed how the evaluation function complies with the requirements of the 2021 revised Evaluation Policy. This policy presents, for the first time, a comprehensive framework through which self-evaluation (conducted under the aegis of Management) and independent evaluation will be planned, conducted and used. The revised policy seeks to promote complementarity and synergy between the two. The principles underlying the policy include usefulness, impartiality and credibility, transparency, partnership, consultation and collaboration, evaluability, and value for money/cost-effectiveness. The policy also describes the procedures for selection, appointment, dismissal and performance assessment of the IOE Director.

The review has been formative and forward-looking, anticipating how the evaluation function can be strengthened in the future. The panel has taken an impartial view, based on knowledge of best practice, with respect to ethical norms and confidential information.

The methodology included the following elements:

- Review of the 2019 EPR and the implementation of the recommendations.
- Review of IOE's self-assessment report (December 2025) which consists of a summary of IOE's assessment of its functions as per the six key clusters guided by the evaluation questions. At the end of each cluster is a SWOT analysis. The document also contains an update on the implementation of the 2019 EPR recommendations. It focuses specifically on independent evaluation in IFAD and will accompany a separate self-assessment developed by IFAD Management on self-evaluation.
- Review of the report on self-assessment of IFAD's self-evaluation function (draft 16 January 2026). Covering the period 2019–2025, it examines how the self-evaluation system has evolved since the 2019 EPR, how it has been shaped by successive institutional reforms, and how it now supports accountability, learning and performance improvement across the full operational cycle, from design through supervision and completion.
- Review of relevant documents related to the evaluation function such as the evaluation policy, the evaluation manual, guidelines for project completion reports, strategic plans, work programmes and annual reports.
- Systematic assessment of a sample of IOE corporate and country evaluation reports published in the review period. The objective was to provide an overview of the methodologies and overall quality of a sample of IOE evaluation products. A sample of 20 evaluations was selected which includes all corporate-level, thematic, project cluster and subregional evaluations, as well as a purposive selection of nine CSPEs, randomly selected with the support of Copilot AI. Evaluations were assessed against a set of indicators grouped into seven quality dimensions: scope, reliability, construct validity, internal validity, external validity, data analysis validity, consistency.

- Assessment of a sample of self-evaluation reports produced in the review period.
- Review of relevant corporate IFAD documents.
- Survey of IFAD staff and members of the Executive Board to assess familiarity with, assessment of and use of the knowledge from the evaluation function. The survey was prepared in November 2025. To ensure that the survey was well adapted and intelligible from a beneficiary's point of view, the survey was then tested and reviewed by IFAD. It was open for two and a half weeks in November/December 2025 and was closed with 512 views. In total, the survey received 400 individual responses: 280 complete answers as well as 125 incomplete. Among the respondents, eight were members of the IFAD Executive Board, and the rest were from IFAD staff. In sum, the response rate for board members was 22. 2 per cent (8 respondents out of 36), and the response rate for IFAD staff was 43.5 per cent (392 respondents out of 900 IFAD staff).
- Interviews/focus groups in-person and virtual with stakeholders (IOE, the Executive Board, the President and other members of senior Management, country directors, country partners, selected staff, and other key informants).

Appendix 3: List of interviewees

IOE staff

Hannah Den Boer, Evaluation Analyst
 Mona Fetouh, Deputy Director
 Massiel Jimenez, Evaluation Officer
 Steven Jonckheere, Senior Evaluation Officer
 Hansdeep Khaira, Senior Evaluation Officer
 Diane Abi Khalil, Senior Evaluation Officer
 Mikal Khan, Evaluation Officer
 Kouessi Maximin Kodjo, Lead Evaluation Officer and Chief of Section
 Monica Lomena-Gelis, Senior Evaluation Officer
 Anne Sophie Lenoir, Evaluation Assistant
 Sarah Mahdar, Temporary Administrative Assistant
 Indran Naidoo, Director IOE
 Oanh Nguyen, Lead Evaluation Officer and Chief of section Paolo Silveri, Lead Evaluation Officer and Chief of section
 Shaun Ryan, Administrative Associate
 Laure Vidaud, Administrative Associate
 Alexander Voccia, Senior Evaluation Communication & Knowledge Management Specialist
 | Coordinator, Evaluation Communication Unit

IFAD staff

Carola Alvarez, Managing Director, Office of Development Effectiveness
 Oscar Anaadumba, MEL specialist
 Stella Atela, Director, People and Culture Division
 Madina Bazarova, Director, Office of Audit and Oversight
 Thomas Bousios, Director, Information, Communication and Technology Division
 Nigel Brett, Director, Sustainable Production, Markets and Institutions Division
 Donal Brown, Associate Vice President, Department for Country Operations
 Han Ulac Demirag, Chief, ODE (Quality Assurance and SECAP Unit)
 Chitra Deshpande, Results specialist, ODE
 Federica Diamanti, Associate Vice-President, External Relations Department
 Roger Fischer, Chief Risk Officer, Office of Enterprise Risk Management
 Lenyara Fundukova, Senior Knowledge Management Specialist, ODE (Evidence for Development Impact and Innovation)
 Ron Thomas Hartman, Director, Global Engagement, Partnerships and Resource Mobilization

Bernard Hien, Director, West and Central Africa Division

Carlos Manuel Icaza Lara, Specialist Policy and Results, ODE (Results, Resources and Systems Unit)

Alvaro Lario, IFAD President

Sara Mbago-Bhunu, Director, East and Southern Africa Division

Rocio Medina Bolivar, Director, Latin America and the Caribbean Division

Vibhuti Mendiratta, Senior Economist, ODE (Evidence for Development Impact and Innovation)

Juan Carlos Mendoza, Director, Environment, Climate, Gender and Social Inclusion Division

Diane Menville, Associate Vice-President, Financial Operations Department

Gerardine Mukeshimana, Vice-President

Robert Palmer, delegated by Jenny Scharrer, Director, Private Sector Operations Division

Tisorn Songsermsawas, Senior Economist, ODE (Evidence for Development Impact and Innovation)

Reehana Rifat Raza, Director, Asia and the Pacific Division

Naoufel Telahigue, Director, Near East, North Africa, Europe and Central Asia Division

Claudia Ten Have, Secretary of IFAD

Leon Williams, Senior Specialist (M&E), ODE (Results, Resources and Systems Unit)

Guoqi Wu, former Associate Vice President, Corporate Services Department

Executive Board Representatives

Ms Lorena Alvarqado Quezada, Alternate Permanent Representative, Permanent Mission of Mexico to the United Nations Organizations for Food and Agriculture in Rome

Mr Jan Bade, Chargé d'Affaires a.i., Permanent Representation of the Kingdom of the Netherlands to the United Nations Organizations for Food and Agriculture in Rome

Mr Nicolas Cattaneo, Collaborateur diplomatique, Embassy of the Swiss Confederation

Mr James Catto, Director, Office of International Development Policy, Department of the Treasury of the United States of America

Mr Edoardo Ciammariconi, Communication Officer and Assistant to the Ambassador, Swiss Mission to the United Nations Agencies in Rome

Mr Miguel Garcia Winder, Former Mexico Executive Board Representative

Ms Marlene Heeb, Deputy Permanent Representative, Counsellor, Swiss Mission to the United Nations Agencies in Rome

Ms Aliisa Ketonen, Advisor, Embassy of the Republic of Finland

Ms Léa Kvaternik-Borderon, Alternate Permanent Representative, Permanent Representation of France to the United Nations Agencies in Rome

Mr Ronald Meyer, Former Germany Executive Board Representative

Ms Susann Nilsson, Executive Board Representative of Sweden to WFP and IFAD

Dr Yaya Olaniran, Minister, Permanent Representative of the Federal Republic of Nigeria to the United Nations Food and Agriculture Agencies in Rome and Evaluation Committee Chairperson

Mr Mohamed Sawy, Executive Board Representative of the Arab Republic of Egypt to the United Nations Food and Agriculture Organizations in Rome

Ms Divyadharshini SHANMUGAM, Director (OMI), Department of Economic Affairs, Ministry of Finance of the Republic of India

Ms Sui Li, Deputy Permanent Representative of People's Republic of China to the United Nations Agencies for Food and Agriculture in Rome

Ms Anja Wagner, Executive Board Representative of the Federal Republic of Germany to the International Organizations in Rome

Country-level Stakeholders

Mr Daniel Chakwesha, International Cooperation, Ministry of Finance, Economic Development and Investment Promotion. Harare, Zimbabwe

Ms. Tu Duong, Administrative Director, RYNAN SMART AGRICULTURE JSC, Viet Nam

Dr Mohamed Yammai, Director General, Resource Mobilization, Resource Mobilization Department, Ministry of Finance and Planning, Male', Maldives

Other

Ms Anne-Claire Luzot, Director, Office of Evaluation, World Food Programme

Mr Fabrizio Felloni, Deputy Director, Independent Evaluation Office, Global Environment Facility

Dr Srini Pillay, Chief Medical Officer and co-founder of Reulay

Dr Vinod Thomas, Senior Adviser for IOE

Rob D. van den Berg, member for IOE's External Advisory Panel 2022-2024

Appendix 4: Surveys of staff and members of the Executive Board

The survey was launched on 26 November 2025 and disseminated by email to IFAD staff by the IOE team. It remained open for two and a half weeks and closed on 12 December 2025, which aligns with good practice for internal organizational surveys. The survey was designed to be short, with an average completion time of 10 minutes, balancing sufficient participation time with respondent engagement. However, the timing of the survey, towards the end of the calendar year, may have affected participation rates, particularly among senior staff and Executive Board members.

The survey was closed with 512 views. In total, the survey received 400 individual responses: 280 complete answers, as well as 125 incomplete. Among the respondents, eight respondents were members of the IFAD Executive Board, and the rest were regular IFAD staff. In sum, the **response rate for Board members was 22.2 per cent** (8 respondents out of 36), and **the response rate for IFAD staff was 43.5 per cent** (392 respondents out of 900 IFAD staff).¹⁷ About 20 per cent of the respondents from IFAD staff were not familiar with the evaluation work conducted in the organization and thus were taken out of the survey. The **final sample for the rest of the survey is 272 IFAD staff and 8 members of the Executive Board**.

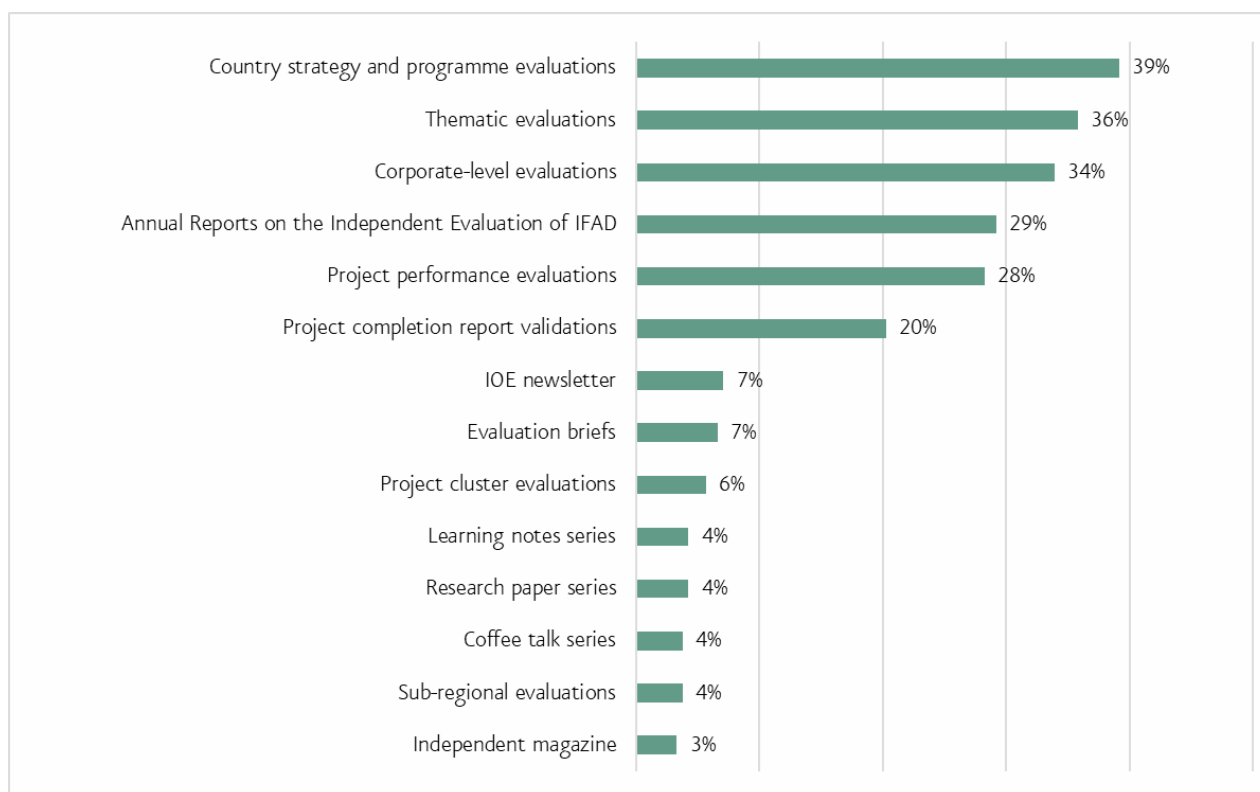
The survey successfully reached a diverse set of IFAD staff:

- Most **respondents are professional staff (54 per cent), followed by consultants (25 per cent)** and administrative staff (15 per cent). Only 3 per cent hold positions at the level of Deputy Director or above. Notably, three respondents identified as interns and two as United Nations volunteers.
- Most of the **respondents (54 per cent) work in the headquarters of IFAD in Rome**, and over a quarter (27 per cent) work in country offices. The next largest group at 12 per cent work in regional offices and, more marginally, 2 per cent (6 respondents) worked remotely, 2 per cent as consultants (5), 1 per cent in a multi-country office (4) and in a liaison office (3). Over a third (34 per cent) of respondents work in the Country Operations Division, followed by 13 per cent each in the Corporate Services Division and the Office of Technical Delivery. The Independent Office of Evaluation accounts for 9.7 per cent of the sample.
- Nearly **half of respondents (47 per cent) have worked at IFAD for one to five years**, while a quarter (25 per cent) has been with the organization for over a decade. Thirteen per cent have less than one year of service and 15 per cent have worked at IFAD for six to ten years.
- The survey sample is **48 per cent female and 46 per cent male**, with 5 per cent of respondents choosing not to disclose their gender.

While the survey reached a diverse range of staff in terms of division, tenure and occupation, limited background data on the overall staff population prevented a full assessment of sample representativeness. Although the survey targeted permanent IFAD staff, responses to screening questions suggest that some non-permanent staff also participated. As a result, the findings should not be considered representative of, or generalized to, the broader population of consultants working with IFAD.

¹⁷ The IFAD staff population figures were supplied by IFAD Human Resources and correspond to data as of October 2025. From our experience, these response rates fall within acceptable ranges (30-50 per cent) for voluntary organizational surveys conducted in multilateral institutions.

**Figure 2 : Which of the IOE products do you find the most useful for your work?
(n = 212, IFAD staff)**



Overall, the **survey collected evidence on a very positive assessment of the IOE:**

- The independence of IOE is the most strongly endorsed quality of the evaluation function, with 51 per cent of respondents strongly agreeing. Assessments are more mixed with regard to the rigour of IOE approaches and methodologies, as well as IOE's engagement with stakeholders. These areas received the highest levels of negative feedback (between 3 to 8 per cent), although they also recorded the largest share of "don't know" responses.
- The survey also collected **highly positive feedback on the usefulness of IOE evaluation recommendations**. Most respondents (82 per cent) reported that the recommendations were useful, including 39 per cent who rated them as very useful. This is **consistent across occupation status and tenure period** at IFAD among all the staff, but it is particularly the case for administrative staff and the deputy directors and above.
- On the effectiveness of IOE's results and communication, the results are somewhat positive with respectively 62 per cent and 55 per cent of respondents stating that it is effective or very effective. Less than a third of respondents said it was only a little effective.
- The assessment of the impact of IOE's work is also positive, with between 65 per cent and 74 per cent of respondents providing positive ratings. However, it should be noted that a relatively high share of respondents, up to 20 per cent, were unable to assess these elements, while negative responses accounted for between 10 per cent and 15 per cent.

The negative feedback regarding IOE mainly relates to the significant workload required from staff during evaluations, the timing of evaluations not always aligning with operational needs and a perceived distance between operational teams and IOE's independent role, as well as concerns that some recommendations are too general or theoretical and not sufficiently adapted to specific operational contexts.

"In my previous role as CPM, after CSPE I tried to involve IOE colleagues at the early stage of a programme design but the participation level or the notion was IOE is "independent" and (we) therefore didn't get much value or input at early stage of the design but rather at a later stage to check on compliance which was late in my view."

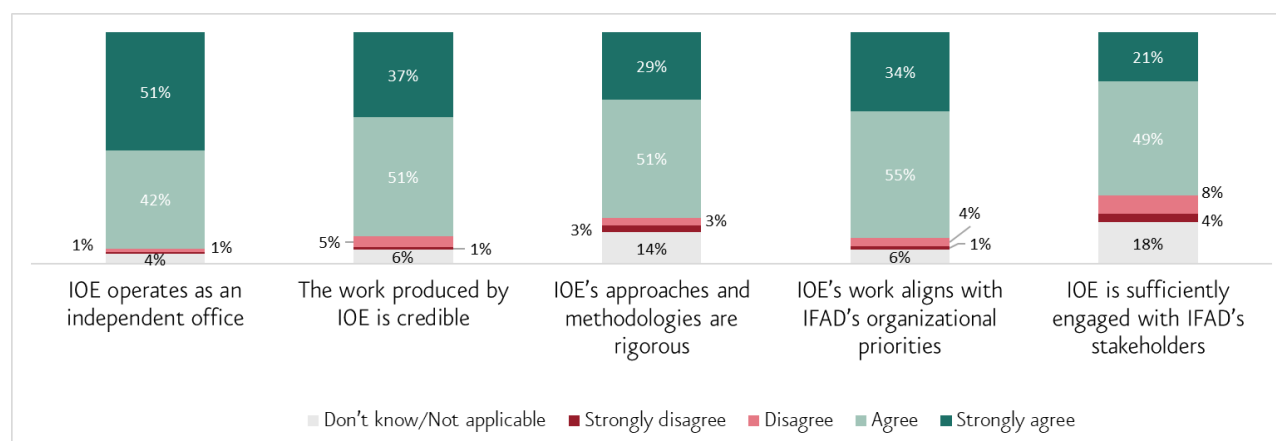
The recommendations call for broadening the scope of evaluations beyond IFAD operations, improving communication through shorter and more accessible outputs and easier access to report findings, strengthening dialogue between evaluation functions and Management, and further decentralizing evaluation functions to bring them closer to country and regional contexts. A selection of survey respondents' comments is illustrative.

"IOE products are somewhat hard to find. They should have more presence in the intranet and more diagrams for easy digest of information."

"In my opinion, staff members and other stakeholders require assessments to be presented in two ways. There are times when in-depth rigorous research is required, of course, but I would also contend that there is a time when 'lighter' briefs, learning notes, and shorter reports are needed, especially when staff are busy. These give readers an understanding of a particular issue, keeping them updated and informed; if they need more information, they can always turn to more in-depth work."

"Given the organizational structure in terms of the independence of IOE, its role in facilitating lessons learning or uptake of evaluations particularly at the regional and country level is somewhat limited. This would work better if the organization had a decentralized evaluation function which has a dotted line management to IOE. This perhaps requires either partial decentralization of IOE in line with what other UN agencies have done or strengthening the self-evaluation function to enhance uptake."

Figure 3 : Based on your experience, how much do you agree with the following statements about these aspects of the IOE? (n = 213, IFAD staff)

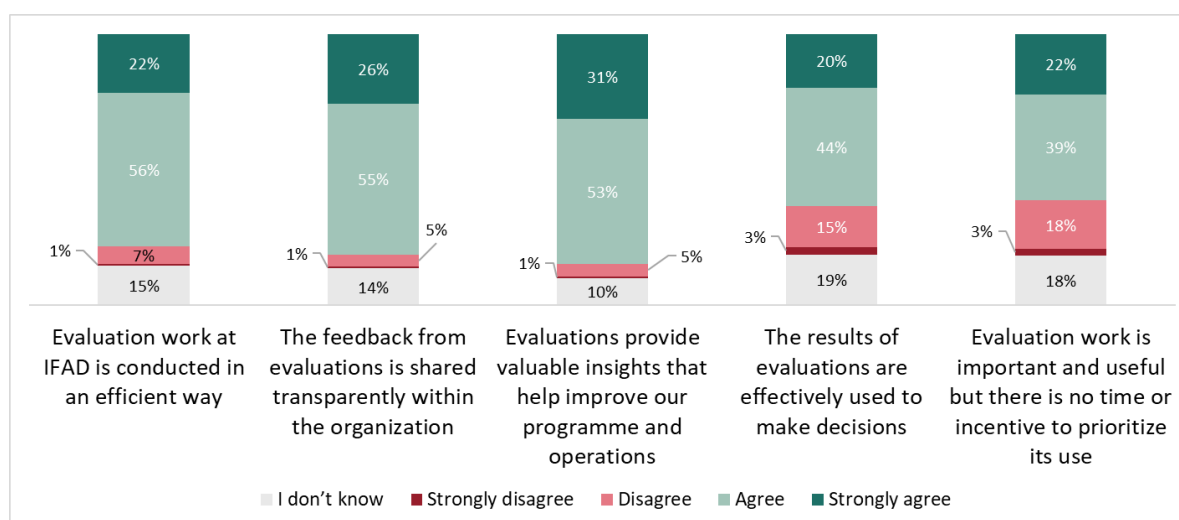


Concerning **Management-led self-evaluation**, less than one third of respondents (31 per cent) reported having been involved in a project completion report (PCR), mostly consultants and professional staff. However, for those concerned with PCRs, their reported use was high, with 39 per cent indicating they use them very often and 51 per cent sometimes; only 5 per cent reported never using them. PCRs are very positively assessed by respondents, with overwhelming majorities (88–97 per cent) agreeing that they support the assessment of achievements, outcomes and impacts, facilitate experience-sharing and lesson-learning, and contribute to improvements in future project design and implementation.

The last section of the survey was dedicated to a **broader assessment of the evaluation conducted** at IFAD:

- Overall, a large majority of respondents express positive views on evaluation at IFAD (with 75–84 per cent positive agreement across statements), particularly regarding the usefulness of evaluations for improving programmes and informing decisions, although a notable minority highlights constraints related to time and incentives for prioritizing their use.
- The effective use of evaluation results in decision-making is slightly less positively assessed with 18 per cent of the respondents disagreeing with this statement.
- The lack of time and incentive to prioritize the use of evaluation also shows a higher level of disagreement among respondents (21 per cent).
- Most respondents (90 per cent) consider IFAD to be a learning organization, including 37 per cent who strongly affirm this view. This perception is generally consistent across tenure and slightly more positive among staff with over six years of service, although administrative staff are somewhat less positive, with 15 per cent stating that this is not at all the case.

Figure 4 : Based on your experience, how much do you agree with the following statements about evaluation work conducted at IFAD (including self-evaluation and independent evaluation)? (n = 197, IFAD staff)



- The most frequently cited enablers of learning are collaboration and teamwork across projects and departments (71 per cent) and access to training, resources and development programmes (58 per cent), followed by feedback mechanisms (41 per cent), with additional factors including professional interest, internal communication and learning events, peer learning and IOE events and newsletters.
- The main barriers to learning identified by respondents are limited resources (71 per cent) and competing priorities (60 per cent), followed by ineffective knowledge management (42 per cent) and hierarchical rigidity (40 per cent), while fewer respondents cite unclear learning objectives (25 per cent), and additional constraints include challenges in headquarters and country office relations regarding decentralized knowledge management and potential tensions with political priorities.

II. Main results from the Executive Board

As described above, the survey collected the complete responses of seven members of IFAD's Executive Board. Six of them are from List A countries, one of them is from List B.

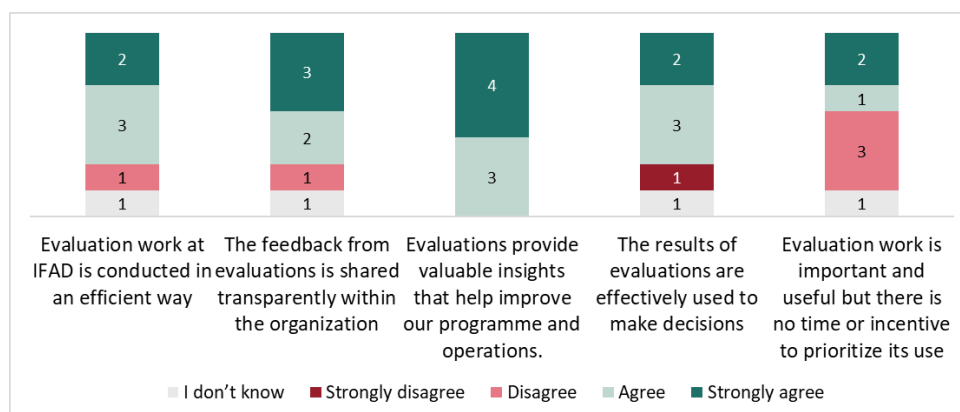
We had no answers from list C countries. This distribution suggests that IFAD's evaluation function may face challenges in engaging Executive Board representatives from List B and List C countries. Almost all of the board member (86 per cent) has read some IOE products, but only about half of them (43 per cent) have used them to be informed about their work.

"I mainly read evaluations. I found them to be very good, balanced and realistic: I like that they aim to facilitate learning and thus further improve IFAD's work."

There is broad agreement that IOE effectively contributes to improving the design and implementation of future projects, with most Executive Board members viewing IFAD as either fully or partially a learning organization. Executive Board members generally find IOE's recommendations useful (57 per cent agree they are very useful) and recognize its independence, credibility, methodological rigour and alignment with IFAD's priorities. Most also view IOE as effective in communicating results, sharing experiences and promoting lesson learning, though some note limited stakeholder engagement and familiarity with outreach to partner countries. One board member expressed some concern:

"IOE talk a lot about themselves but less about the outcome or results of their work."

Figure 5 : Board member assessment of IOE (n = 7, Executive Board)



Leadership support, a learning-friendly culture, collaboration and access to training are seen as the main enablers of learning at IFAD, though feedback mechanisms and the knowledge management system are less recognized, with some caution about problem cover-ups and external pressures. Key barriers include competing priorities, limited time and resources, ineffective knowledge management, hierarchical rigidity, unclear learning goals, and, for some, pride or protection of financial resources.

"There seems to be more cover-up if problems arise than interest in learning - this is a huge worry for us."

Appendix 5: Quality review of IOE evaluation reports

This meta-evaluation is based on a review of IOE evaluations conducted between 2019 and 2025. The quality review was prepared in January 2026, and the report was drafted in early February 2025. The analysis and report were produced by Matthieu Sarnin, Consultant at Quadrant Conseil.

The objective of this review was to provide an overview of the methodologies and overall quality of a sample of IOE evaluation products, in order to inform the independent external review of IOE. While an in-depth review of individual evaluations was beyond the scope of this exercise, the review sought to identify recurring strengths and weaknesses across evaluations conducted by IFAD.

The review universe comprises 40 evaluation reports, including five corporate-level evaluations (CLEs), two thematic evaluations (TEs), two project cluster evaluations (PCEs), two subregional evaluations (SREs), and 29 country strategy and programme evaluations (CSPEs). From this universe, a sample of 20 evaluations was selected. The sample includes all corporate-level, thematic, project cluster, and subregional evaluations, as well as a purposive selection of nine CSPEs, randomly selected with the support of Copilot AI. The complete list of reviewed reports is presented in annex I.

The systematic quality review of IOE evaluations applied a simplified version of the meta-evaluation framework developed by the Independent Evaluation Group (IEG).¹⁸ Evaluations were assessed against a set of indicators grouped into seven quality dimensions: scope, reliability, construct validity, internal validity, external validity, data analysis validity, consistency. For each of these attributes, we assess a series of indicators based on a grading scale (1) adequate, (2) partial, (3) inadequate, or inapplicable (NA). The key questions of the framework can be found in annex II. Based on the findings, the review highlights several key quality concerns and identifies areas requiring further attention.

I. Limitations

Finally, several limitations affecting the quality review should be noted. First, contrary to the IEG meta-evaluation framework, the analysis relies exclusively on final evaluation reports published by IFAD¹⁹ and comparison with methodology presented in approach papers was not possible. In addition, the IEG framework includes a second phase consisting of an in-depth quality review of a limited subsample of evaluations, complemented by interviews with evaluation staff to capture contextual information on evaluation design and implementation. This level of qualitative depth was beyond the scope of the present study. Given time constraints and the larger number of evaluations reviewed, a more systematic but less in-depth approach was applied to collect information on evaluation. Secondly, the scope and nature of the evaluations included in the review vary considerably. However, evaluations were treated as equivalent types of products during the data collection phase, meaning that the same items were reviewed for each evaluation, the diversity of evaluation types and their respective idiosyncrasies were considered during the analysis phase and should also be taken into account when interpreting the results presented in this report.

¹⁸ For brevity, this report omits detailed definitions and rationales for each attribute in the framework. The full methodology is available in: World Bank. 2021. A Meta-Evaluation of Independent Evaluation Group Evaluations (Fiscal Years 2015–19). Independent Evaluation Group. Washington, DC: World Bank.

¹⁹ The exercise did not cover dimensions of credibility and quality that could not be assessed based on the reports alone, such as consultations between evaluators and counterparts, the expertise and composition of evaluation teams, integrity in data collection, the nature of quality assurance processes, and the quality of peer review, where applicable.

Considering these limitations, this report should not be seen as a standalone product, but rather as a complementary input intended to inform the external evaluation review.

II. Results

The assessment finds that overall, the quality of the IOE reports is very good and that in most cases, the evaluation reports clearly describe and report the key elements under review, as presented in the score summary in Appendix 5. Accordingly, the challenges highlighted below mainly reflect areas where further improvements could be made from an already strong baseline.

Assessment of the overall evaluation design

The scope of evaluations varies substantially depending on the type of evaluation product. While some evaluations focus on a limited number of programmes (e.g. PCEs) or a single country (e.g. CSPEs), corporate and thematic evaluations tend to adopt a much broader and more ambitious scope. In addition, the timespan covered by the evaluations is often extensive, with half of the sample covering periods of more than ten years. As a result, the evaluands described in the sample are frequently highly complex, reflecting the multisite, multilevel and multi-actor nature of IFAD-supported interventions. Most evaluations in the sample explicitly stated a learning objective with only three which did not refer to learning or lesson drawing as a stated purpose. In addition, the evaluations can broadly be grouped according to how their objectives are framed: one group adopts a more causal perspective, while the other emphasizes performance assessment.²⁰ Finally, regarding evaluation questions, all evaluations except four formulated between two and five overarching questions to guide the evaluation process. None subsequently introduced an excessive number of sub-questions that would resemble a “bag-of-questions” approach to scoping. Most overarching questions were evaluative in nature. To a lesser extent, approximately half of the evaluations included exploratory questions, nine out of fourteen incorporated design-oriented questions, and only two raised more descriptive questions aimed at understanding programme contexts. Given the ambitious scope of many evaluations, this limited emphasis on contextual understanding is somewhat surprising, as it constrains deeper insight into the environments in which IFAD-supported interventions operate.

Despite this, most recommendations were logically grounded in the evidence presented throughout the evaluation reports. One minor caveat is that, although recommendations were generally evidence-based, there is a tendency to accumulate what could be described as “bags of recommendations.” Evaluations typically included four to seven overarching recommendations, but approximately half of the evaluations also specified up to 22 sub-recommendations. While balancing the specificity and breadth of recommendations is inherently challenging, the “bag of recommendations” approach risks diluting the key messages of the evaluation and potentially creating confusion about priorities and focus.

Reliability and quality of evidence issue

All evaluation relied on a well-developed evaluation design matrix and a theory of change, which provides a clear rationale for the methods used to address specific evaluation questions. Findings were generally well described and logically linked to the underlying data and methods. However, one area of concern is that most evaluations do not provide sufficient clarity on how evidence from different sources was triangulated and used to substantiate findings relative to the initial evaluation questions. Several important cautions should be noted regarding the transparency and clarity of the analysis of evidence.

²⁰ This distinction appears largely driven by the type of evaluation, with more cross-sectoral evaluations (e.g. CLEs and TEs) focusing on IFAD's contribution to broader social outcomes, and more narrowly-scoped evaluations (e.g. PCEs and CSPEs) concentrating on project-level results in relation to predefined objectives, often linked to country strategies.

This relates to four elements reviewed: the description of data analysis methods, the explanation of how evaluative evidence was triangulated and synthesized, the articulation of quality assurance principles applied during coding and synthesis, and the depth of the discussion of limitations. Only two evaluations explicitly described how data were analyzed and how evidence was triangulated. Similarly, only two evaluations included elements related to analysis and triangulation within the evaluation design matrix itself. The evaluation titled Joint evaluation of collaboration among the United Nations Rome-Based Agencies (ID3CLE)²¹ offers a good example of a clearly-formulated data analysis approach: “Data analysis was structured around the questions, sub-questions and indicators in the evaluation matrix. The evaluation team developed a standardized findings matrix template, structured by evaluation questions and sub-questions. Team members entered data from each of the country, deep dive and thematic studies into one of these matrices, and all matrices were then combined for easy analysis. Structuring the data in this way facilitated triangulation. Analysis of documentation and review of interview notes supplemented the combined findings matrices. The evaluation team integrated its understanding and analysis from all these sources through a series of workshop discussions. [...]” (ID3CLE). Conversely, the evaluation Knowledge Management Practices in IFAD Corporate-Level Evaluation (ID1CLE) provides a useful example of the “black box” approach to analysis and triangulation often found in the evaluations: “During the synthesis phase, the evaluation team assessed the extent to which the evidence was adequate and could be triangulated through the various sources and methods. Constraints in the availability and quality of data were, in particular, noted for evidence blocks 1 and 2 (see limitations).” (ID1CLE, own emphasis)

Second, a closely-related issue identified in the systematic review concerns the quality assurance principles applied during coding and synthesis. All evaluations but one included in the sample underwent an internal peer review process within IOE. Eight evaluations reported the involvement of external independent consultants in quality assurance and enhancement activities, and half explicitly referred to compliance with IOE evaluation standards or internal guidance. However, only a limited number of evaluations provided detailed information on the quality assurance principles applied throughout the evaluation process, the coding of data and the analysis.

Finally, a well-developed discussion of limitations is an important factor affecting the reliability of an evaluation. In this respect, most evaluations performed somewhat well, as all included a dedicated section or paragraph outlining key limitations and, in most cases, the mitigation approaches adopted. However, while limitations on data sources and methods are generally well documented, only three evaluations addressed limitations in the evaluation design and reflected on the evaluation process. Finally, the extent to which mitigation measures were discussed in response to identified limitations varies considerably across the evaluations analyzed. Some evaluations clearly explained how specific methods were used to address evidence gaps. This is the case, for example, in the country strategy and programme evaluation (CSPE) of the People’s Republic of China: “To ensure adequate coverage of evidence, qualitative interviews, an extensive literature review and an online stakeholder survey complemented this CSPE to the extent possible and allowed triangulation of quantitative and qualitative information.” (ID13CSPE). By contrast, other evaluations remained vague regarding the mitigation measures undertaken. The CSPE of the Republic of Türkiye illustrates this limitation: “In light of these limitations, the team triangulated using diverse sources of data and information, where accessible, before concluding.” (ID12CSPE).

²¹ It should be noted that this evaluation is the only one in the sample conducted by an external private company, rather than internal evaluators. While the emphasis on transparency and clarity in source analysis is understandable in this context, internally-conducted evaluations should also adhere to similar standards.

A final concern regarding the reliability of evaluation results relates to the extent to which findings can be expected to apply in other settings. Overall, temporal validity was generally well addressed in the evaluation reports, with evaluations reflecting on the maturity of projects and the timing of data collection. In most cases, external validity was discussed in relation to the selection of primary data sources, such as portfolio analyses, country case studies and field visits, and was also considered when assessing secondary data. For example, in the Republic of Türkiye country strategy and programme evaluation (ID12CSPE): “In the AKADP, the samples used in the impact assessment survey had no links to those in the baseline survey. In the MRWDP, the Government of Türkiye impact assessment did not use the baseline data. The validity of its comparison of results between beneficiaries and control groups is questionable without an understanding of these two groups before the project”. However, discussions of external validity specifically regarding the generalizability of evaluation findings were relatively scarce and, in some cases, omitted entirely. Population validity was particularly weak in most reports, leading to partial ratings for several evaluations. Some evaluations, however, acknowledged limitations related to external validity in a clear and meaningful way. For example, the Sub-Regional Evaluation of Countries with Fragile Situations in IFAD-WCA, which examined IFAD’s engagement in the G5 Sahel countries and northern Nigeria, explicitly addressed generalization issues arising from limited data sources.

Overall, while the evaluations reviewed in the sample were generally robust with respect to data analysis validity, concerns related to data quality and the strategies used to mitigate potential biases arising from weaker data emerged as key areas for improvement under this attribute. The main issue raised by these findings is that, although evaluations collect large volumes of evidence from a wide range of data sources and methods, it is often difficult to assess the relative strengths and weaknesses of that evidence, which tends to be reported in a largely linear and non-hierarchical manner that can lack analytical clarity. A more explicitly articulated link between evidence, methods and the scope of inference would improve the interpretability and robustness of the findings. This could be done through a tighter articulation of the theory of change in relation to the empirical components of the study and could enhance transparency in how evidence is synthesized and weighted, thereby improving the overall rigour of the evaluation.

IFAD-centred approach

All the evaluation reports included in the analysis used a theory-based approach, reconstructing a theory of change, sometimes situated within a broader conceptual framework and they often provided a compelling section to operationalize key concepts within the formulation of the theory or the conceptual framework used. However, only half of the evaluations mobilized and incorporated external research following structured review principles. Consequently, most evaluations tend to adopt an IFAD-centred perspective in defining and operationalizing key concepts and fall short in describing the broader policy context or the roles of external actors engaged alongside IFAD interventions. A notable exception is the corporate-level evaluation of Knowledge Management Practices in IFAD (ID1CLE), which illustrates strong use of external literature. In this case, the conceptual framework and nested theory of change draw on an extensive review of the six generations of knowledge management functions to explore “a set of hypotheses that cover causal linkages between KM strategies, tools, practices and key enabling factors, and how they interact to deliver KM results at corporate, regional division and country programme levels” (ID1CLE). The assessment of construct validity thus highlights that including external knowledge more systematically and applying structured literature review methods could broaden the evaluation perspective beyond IFAD-specific interventions.

III. Conclusion

Overall, the evaluations reviewed demonstrate high quality and reflect a strong appreciation of methods and evidence, even in the context of ambitious evaluation scopes and complex evaluands. At the same time, several areas for improvement were identified. These include greater transparency regarding data quality and triangulation, more systematic attention to the environment in which IFAD operates, and enhanced reflection on the external validity of findings. Strengthening the use of the theory of change, particularly its articulation in relation to empirical analysis, would further improve analytical clarity. More broadly, adopting a more rational and formalized approach to evaluation design, moving beyond a collection of methods and tools toward a coherent, structured framework, would support a deeper understanding of evaluation quality. The integration of innovative approaches, such as contribution analysis, could also enhance assessments of coherence and causal inference. Expanding attention to these dimensions would improve the validity, reliability and overall utility of future evaluations, while maintaining the high standards of methodological rigour already evident across the sample analyzed.

Annex I : List of evaluations in the review sample

ID	Title
ID1CLE	Corporate-level evaluation of knowledge management practices in IFAD. December 2024
ID2CLE	Corporate-level evaluation of IFAD's decentralization experience 2023. November 2023
ID3CLE	Joint evaluation on the collaboration among the United Nations Rome-based agencies. October 2021
ID4CLE	Corporate-level evaluation on IFAD's support to innovations for inclusive and sustainable smallholder agriculture. September 2020
ID5CLE	Corporate-level evaluation on IFAD's engagement in pro-poor value chain development. September 2019
ID6TE	Thematic evaluation of IFAD's support for smallholder farmers' adaptation to climate change. May 2023
ID7TE	Thematic evaluation of IFAD's support to gender equality and empowerment. March 2025
ID8PCE	Project cluster evaluation on rural finance in the East and Southern Africa Region. December 2023
ID9PCE	Project cluster evaluation on rural enterprise development. August 2023
ID10SRE	Sub-regional evaluation of IFAD's experience in the Dry Corridor of Central America. August 2024
ID11SRE	Sub-regional evaluation of countries with fragile situations in IFAD-WCA. Learning from experiences of IFAD's engagement in the G5 Sahel countries and northern Nigeria. March 2023
ID12CSPE	Türkiye country strategy and programme evaluation. 23 April 2025
ID13CSPE	China country strategy and programme evaluation. 21 February 2024
ID14CSPE	Colombia country strategy and programme evaluation. 7 February 2024
ID15CSPE	Malawi country strategy and programme evaluation. 24 January 2023
ID16CSPE	Islamic Republic of Pakistan country strategy and programme evaluation. 4 March 2022
ID17CSPE	Uganda country strategy and programme evaluation. 7 July 2021
ID18CSPE	Ecuador country strategy and programme evaluation. 3 March 2021
ID19CSPE	Madagascar country strategy and programme evaluation. 1 December 2020
ID20CSPE	Tunisia country strategy and programme evaluation. 8 March 2019

Annex II : Quality review assessment

Attribute	Review
Scope	- Are the context and rationale of the evaluation adequately described? (Graded) - Are the evaluation goals adequately formulated? (Graded) - Are the evaluation questions adequately formulated (also in relation to each other)? (Graded) - Are the evaluation questions adequately linked to the evaluation goals? (Graded) - Has the scope and delimitation of the evaluation been adequately described? (Graded)
Reliability	- Is the methodology of the evaluation adequately described? (Graded) - Is there a theory of change or theory of action, or conceptual framework? (Y/N question) - Is there identification and analysis of an operations portfolio? (Y/N question) - Are the quality assurance principles in coding and synthesis described? (Y/N question) - Are the sampling and selection considerations described? (Y/N question) - Is there a description of data collection methods and sources of data? (Y/N question) - Is there a description of data analysis methods? (Y/N question) - Is there a triangulation and synthesis of findings, including how (different) findings from different methods or designs have been integrated to reach (general) conclusions? (Y/N question) - Are the limitations of the evaluation adequately described (resulting from limitations in scope, methods or data and validity of findings)? (Y/N question)
Construct validity	- Has the evaluation clearly defined and operationalized key concepts into measurable attributes, and are these concepts adequately interrelated within a coherent theoretical framework (theories of action, change and conceptual models)? (Graded) - Has the evaluation effectively utilized and integrated external literature following structured review principles, and if applicable, has it successfully implemented a theory-driven approach with a clear articulation of mechanisms, contexts and outcomes? (Graded)
Internal validity	- How does the evaluation address causality and attribution? (If there are no causality questions, this does not apply. If there are causality questions, more information is needed to grade the dimension) (Graded) - In the presence of causal questions: - Is the identification strategy based on theory or counterfactual analysis, or is there an evident triangulation of different sources? (in this particular question, you must extract the identification strategy defined in the evaluation).
External validity	- Has the report paid adequate attention to population validity or generalizability of findings (the ability to generalize the study results to individuals or target groups, organizations or regions not included in the study)? (Graded)
Data analysis validity	- Has the evaluation addressed risks of bias resulting from data weaknesses or demonstrated any signs of incorrect methodological application in the qualitative or quantitative analyses? (Graded)

Consistency	- Are the methods and data sources logically linked to the evaluation questions? (Graded) - Have the methods that are reported as being applied indeed been applied? (Graded) - Do the findings logically relate to the underlying data and methods used? (Graded) - Do the findings respond to the original evaluation questions? (Graded) - Do the recommendations logically flow from the findings? (Graded)
--------------------	---

code name	Scope	Reliability	Construct validity	Internal validity	External validity	Data analysis validity	Consistency
ID1CLE	Adequate	Partial	Adequate	Adequate	Adequate	Adequate	Adequate
ID2CLE	Adequate	Partial	Partial	Partial	Adequate	Partial	Adequate
ID3CLE	Adequate	Partial	Partial	Adequate	Adequate	Adequate	Adequate
ID4CLE	Adequate	Adequate	Adequate	Adequate	Adequate	Adequate	Adequate
ID5CLE	Adequate	Adequate	Adequate	Partial	Partial	Adequate	Adequate
ID6TE	Adequate	Adequate	Adequate	Adequate	Partial	Partial	Adequate
ID7TE	Adequate	Adequate	Adequate	Adequate	Adequate	Adequate	Adequate
ID8POE	Partial	Adequate	Partial	Adequate	Adequate	Adequate	Adequate
ID9POE	Adequate	Adequate	Adequate	Adequate	Adequate	Adequate	Adequate
ID10SRE	Adequate	Adequate	Partial	Adequate	Partial	Adequate	Adequate
ID11SRE	Adequate	Adequate	Adequate	Adequate	Adequate	Adequate	Adequate
ID12CSPE	Adequate	Partial	Adequate	Adequate	Partial	Adequate	Adequate
ID13CSPE	Adequate	Adequate	Partial	Partial	Adequate	Partial	Partial
ID14CSPE	Adequate	Partial	Partial	Adequate	Partial	Adequate	Adequate
ID15CSPE	Adequate	Adequate	Adequate	Partial	Partial	Adequate	Adequate
ID16CSPE	Partial	Adequate	Adequate	Partial	Adequate	Adequate	Adequate
ID17CSPE	Partial	Partial	Partial	Adequate	Partial	Adequate	Partial
ID18CSPE	Adequate	Adequate	Partial	Partial	Adequate	Adequate	Adequate
ID19CSPE	Partial	Adequate	Partial	Adequate	Adequate	Partial	Adequate
ID20CSPE	Partial	Adequate	Partial	Adequate	Adequate	Adequate	Adequate