

Workshop des AK-Strukturpolitik der DeGEval

29. Juni 2023

Hannover

Kohäsionspolitik:

Einsichten aus Bewertungen von

Evaluierungen in der Programmperiode 2014 -2020

Lydia Greunz

Evaluation Helpdesk u. Applica srl



Präsentation – Überblick

- Evaluation Helpdesk: Hintergrund
- Hauptaufgaben des Evaluation Helpdesks
- Evaluierungstätigkeit in Europa im Überblick
- Einsichten aus Bewertungen von Evaluierungen und Unterstützungsgesprächen mit Verwaltungsbehörden (VB)
 - Allgemeine
 - Leistungsbeschreibungen
 - Auswahl und Anwendung von Methoden
 - Daten
 - Timing
 - Planung

Evaluation Helpdesk: Hintergrund

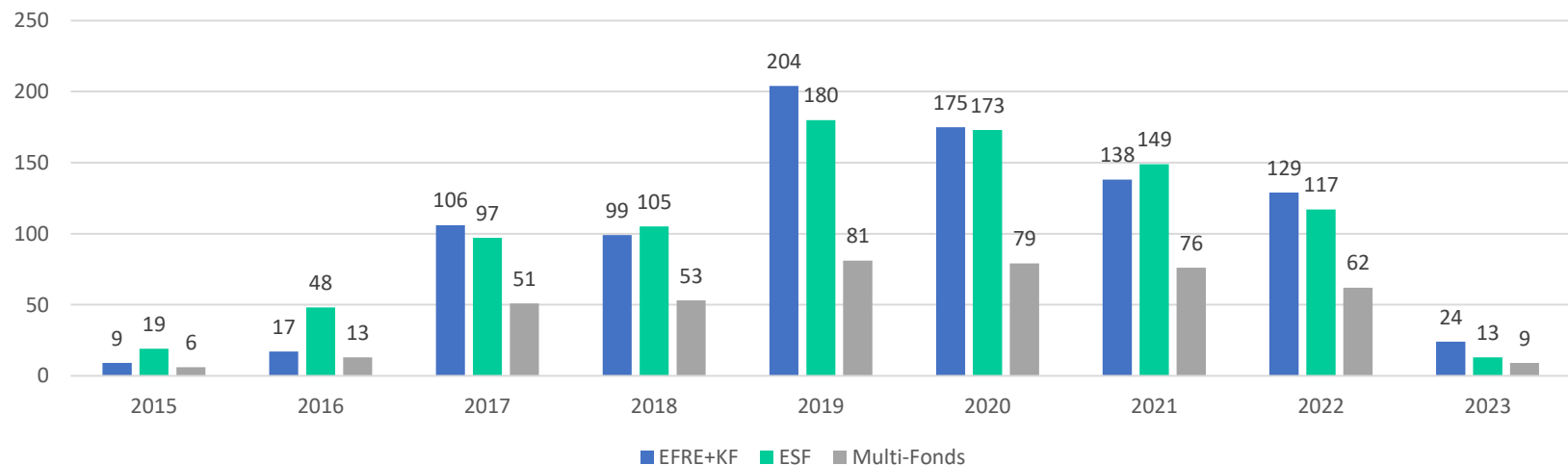
- Evaluation Helpdesk in derzeitiger Form, 2015 eingerichtet um GD REGIO als auch GD Beschäftigung in Evaluierungstätigkeiten zu unterstützen
- Hauptziel ist Verbesserung der Bewertungsqualität der Kohäsionspolitik in der gesamten EU – und letztlich Verbesserung der Programme
- Helpdesk wird von 2 Beratungsunternehmen geleitet: Applica srl mit Sitz in Brüssel und Ismeri Europa mit Sitz in Rom, unterstützt durch internes Team (10 Mitarbeiter) und ein Netzwerk von Beratern und Experten
- Ähnliche, aber weniger umfassende Unterstützung für die Bewertung in GD REGIO seit 2011 – Ad-hoc-Unterstützung und Peer-Review-Aktivitäten in der GD REGIO seit Anfang der 2000er Jahre

Evaluation Helpdesk: 6 Hauptfunktionen

- Bewertung aller Evaluierungen die im Rahmen von EFRE und ESF Programmen in Europa durchgeführt werden – und Veröffentlichung
- Bewertung der Evaluierungspläne von 2014-2020 Programmen in Bezug auf Ziele, Design u. Daten, Organisation u. Fachwissen der Bewertungseinheiten
- Unterstützung von Verwaltungsbehörden (VB) in Evaluierungsfragen einschließlich
 - sachkundige Bewertung von Evaluierungen
 - Beratung zu Leistungsbeschreibungen, Methoden und Daten
 - Beratung zu Bewertungsprozessen, Entwicklung einer Bewertungskultur und Bewertungsmarkt
- Unterstützung von Sachverständigen in GD REGIO in Evaluierungsfragen
- Organisation von „Sommerschulen“ für VB zu spezifischen Themen und Unterstützung der zweijährigen GD REGIO-Evaluierungskonferenz
- Erstellung von „Länderberichten“ und thematischen Berichten und andere Zusammenfassungen der Bewertungsergebnisse

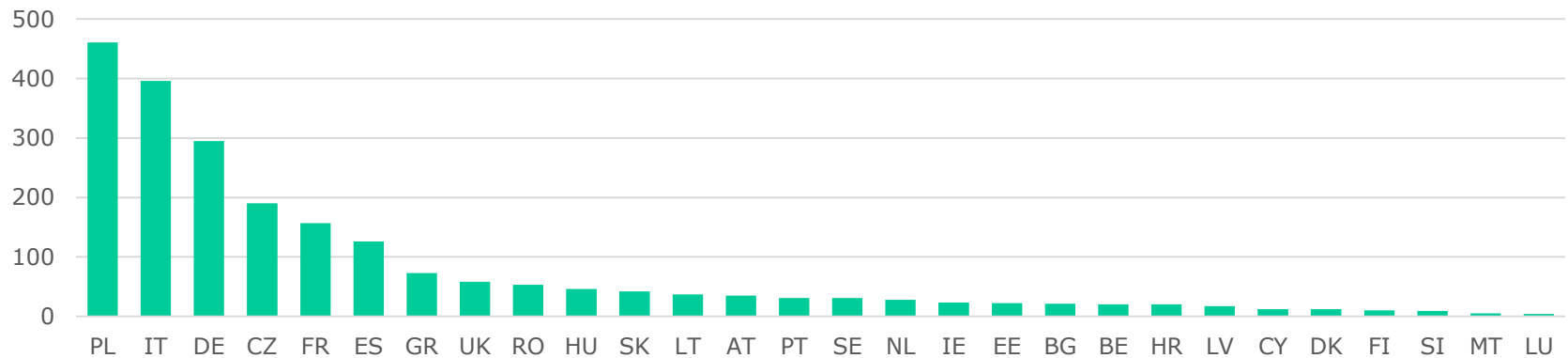
Evaluierungen im Überblick - Stand: Juni 2023

- Seit der Einrichtung des Helpdesks im Jahr 2015 wurden 2,593 Evaluierungen bewertet und zusammengefasst – Zusammenfassungen werden in der Online-Evaluierungsbibliothek der GD REGIO veröffentlicht
- Entwicklung der Bewertungstätigkeit in MS von 2014-2020 Programmen (aufgegliedert nach Fonds und Publikationsdatum)

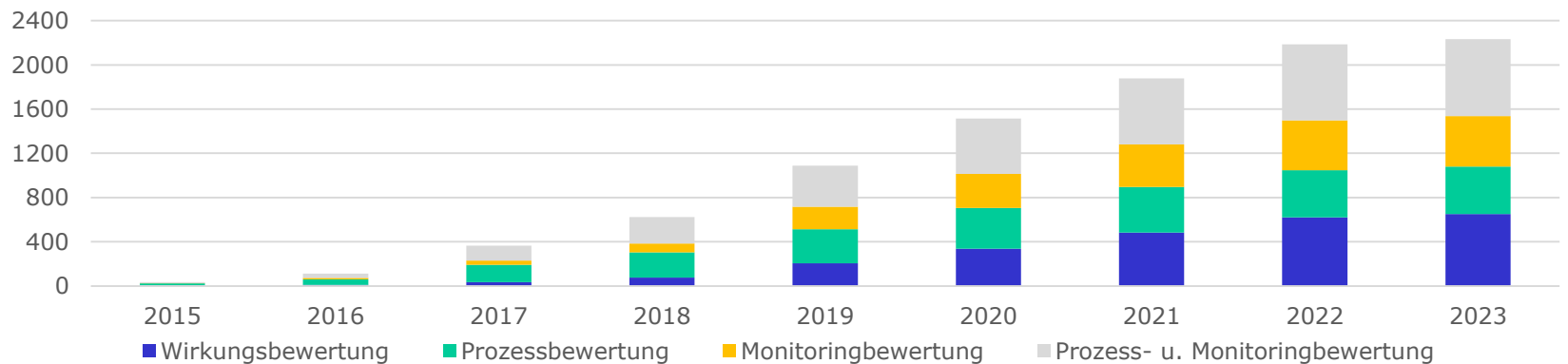


Evaluierungen im Überblick

➤ Evaluierungstätigkeit von 2014-2020 Programmen in MS



➤ Entwicklung des Evaluierungsbestands von 2014-2020 Programmen nach Art



Allgemeine Einsichten

- Aus Besuchen von VB ist klar, dass gute Bewertungen angestrebt werden
- Ziel ist nicht nur Einhaltung der Vorschriften, sondern auch die Bewertung der Wirksamkeit der Programme, der Ergebnisse und der Möglichkeiten zur Verbesserung
- *„Die Bewertung der Kohäsionspolitik ist ein sich wandelndes Bild. Die Evaluierung unterstützt sowohl Rechenschaftspflicht als auch Lernen. Was in Kohäsionspolitik einzigartig ist, ist, wie sich das Verständnis von Rechenschaftspflicht und Lernen entwickelt hat:*
 - *Von ‚Compliance-Checks‘, Validierung und ‚Ausgeben von Budgets‘ bis hin zu Effektivität, Ergebnissen und Umsetzung strategischer Ziele*
 - *Von ‚Brüssel glücklich machen‘ bis zu einem Beitrag zur nationalen und territorialen Politik & Politikentwicklung*
 - *Von der Bewertung, was getan wurde, bis hin zum Lernen, wie man die Dinge besser macht, einschließlich Implementierung und Bereitstellung, um die Ergebnisse zu optimieren.*

*Ungleichmäßige, aber **erkennbare Fortschritte!**“ (Elliot Stern)*

Allgemeine Einsichten

- Aus Besuchen von VB ist auch klar, dass diese Schwierigkeiten haben gute Evaluierungen durchzuführen zum Teil wegen Mängel im lokalen Bewertungsmarkt – mangelnde Kompetenz der Bewerter, die zur Verfügung stehen
- Einige Maßnahmen zur Förderung der Entwicklung und Expansion des Evaluierungsmarktes wären:
 - Förderung des akademischen Engagements + Einsatz internationaler Evaluierungsexperten
 - Unterstützung von Evaluierungsprogrammen oder -kursen an Hochschulen
 - Information der Bewerter über geplante Evaluierungen und Erwartungen
- Die Vergabevorschriften sind auch eine Einschränkung, aber eine allmählich schwächere

Einsichten - Leistungsbeschreibung (LB)

- LB ist „Wurzel“ guter – und schlechter – Evaluierungen
- Gute LB ist wichtig um das zu erhalten, was von einer Evaluierung erwartet wird – Eigenschaften guter LB:
 - klare Angabe der Ziele – was wir wissen wollen und warum – und von Prioritäten, so dass die Bewertung angemessen ausgerichtet werden kann
 - klare Bewertungsfragen und begrenzte Zahl – konzentriert sich auf die wichtigsten Ziele, so dass die Bewertungsbemühungen fokussiert bleiben
 - Fragen, die aufgrund verfügbarer oder zu erhebender Daten beantwortet werden können
 - klare Angabe der Datenverfügbarkeit und der zu füllenden Lücken
 - mögliche Angabe von Methoden - die Bieter sollten Methoden vorschlagen können die sie zur Beantwortung von Fragen zu verwenden beabsichtigen
 - klare Festlegung der Bewertungsanforderungen, einschließlich Fachwissen
 - klare Festlegung der Auswahlkriterien und jeweiliges Gewicht

Einsichten – Methoden Allgemein

- Evaluierungsdesign muss in der Lage sein, Bewertungsfragen zu beantworten
- Bei Wirkungsanalysen ist gewählter Ansatz von entscheidender Bedeutung – muss in der Lage sein, die Wirkung der untersuchten Maßnahmen von anderen Faktoren zu unterscheiden
- Mehrheit der überprüften sogenannten Wirkungsanalysen umfasst lediglich Analyse von ‚Monitoringdaten‘ und Befragungen / Interviews von Unterstützungsempfängern
- Repräsentativität der Befragten, in vielen Fällen nicht berücksichtigt
- Häufig mögliche Verzerrungen von Informationen, die nicht ausdrücklich erkannt oder angepasst wurden – nicht mit Informationen aus anderen Quellen oder mit Verhalten oder Leistung von Nichtempfängern verglichen
- Mangelnde Berücksichtigung möglicher Rückruffehler bei der Angabe der Situation vor Erhalt der Unterstützung
- Umfragen + Interviews können nützliche Beweise liefern, sollten aber idealerweise in verschiedenen Phasen der Intervention durchgeführt werden – insbesondere vorher und nachher

Einsichten - Angewandte Methoden

- Eine Minderheit der Bewertungen verwendet kontrafaktische (KF) und/oder theoriebasierte (TB) Methoden, die darauf ausgerichtet sind, andere Faktoren zu berücksichtigen, die auf Ergebnisse Auswirkungen haben könnten
- Noch weniger Kosten-Nutzen-Analysen – in der Regel für große Projekte
- Das Wesen der KF Bewertungen besteht darin, die Auswirkungen von Interventionen zu identifizieren – beantworten *„ob und wie viel Fragen“*
- TB Analysen versuchen zu identifizieren, wie Effekte erzeugt wurden – um kausale Kette zu identifizieren and zu bewerten, die von einer Intervention zu Ergebnissen führt und die zugrunde liegenden Mechanismen – beantworten *„wie und warum und in welchem Kontext Fragen“*
- Aber für viele Bewertungen ist eine Kombination beider nötig, um zu verstehen, warum und wie Interventionen wirksam sind
- Gute KF Bewertungen sollten mit der Beschreibung der „Wirkungstheorie“ und der Erstellung der Kausalkette beginnen, um sicherzustellen, dass alle relevanten Faktoren berücksichtigt werden, die Auswirkungen auf die Ergebnisse haben könnten

Einsichten - Probleme bei der Anwendung von Methoden

- Bei KF Wirkungsuntersuchungen besteht häufiges Versagen darin, Vergleichsgruppen zu definieren, die den Teilnehmern nicht hinreichend ähnlich sind oder auch die Repräsentativität nicht nachzuweisen
- Die große Schwäche TB Wirkungsanalysen ist die Theorie nicht ausreichend detailliert zu beschreiben, dass Kausalkette + Mechanismen zufriedenstellend untersucht werden können
- In vielen Fällen ist die „Wirkungstheorie“ überhaupt schwer zu erkennen – oft mit Pfeilen dargestellt, die Glieder in einer Kette verbinden, ohne die beteiligten Mechanismen zu erklären
- Das Aufspüren von Nachweisen einer Kausalkette erfordert eine detaillierte Untersuchung durch Feldarbeiten – oder Fallstudien – zur Untersuchung, wie eine Intervention in der Praxis funktioniert. Aber dies wird oft nicht mit ausreichender Anstrengung durchgeführt

Einsichten - Daten

- Der Zugriff auf relevante und zuverlässige Daten ist entscheidend für eine Evaluierung, um robuste Ergebnisse zu erzielen
- Von zentraler Bedeutung ist ein effizientes und vollständiges Monitoringsystem zur Erhebung von Outputs und Ergebnissen
- Aber auch Daten über Entwicklungen, die möglicherweise Auswirkungen auf die Ergebnisse von Maßnahmen haben oder die für kontrafaktische Analysen erforderlich sind, sind ausschlaggebend
- Verwaltungsdaten sind in der Regel die zuverlässigste Quelle, aber nur wenige überprüfte Bewertungen nutzen sie
- Dies nicht weil solche Daten nicht verfügbar sind, sondern weil sie nicht zugänglich sind – in erster Linie aufgrund von Vertraulichkeitsproblemen
- Letztere können häufig durch Anonymisierung von Daten vor der Bereitstellung gelöst werden – erfordern jedoch in der Regel Verhandlungen mit der zuständigen Behörde und die vorausschauende Planung
- „Big Data“ werden auch immer verfügbarer – verwendet zur Bewertung von Verkehrsinvestitionen (in Estland und Polen)

Einsichten - Timing

- Viele überprüfte Wirkungsbewertungen wurden zu früh durchgeführt, um die Auswirkungen von Maßnahmen zuverlässig beurteilen zu können
- Zeit ist allgemeines Thema in Wirkungsanalysen und Zeitpunkt ist selten ideal: entweder zu früh, um die volle Wirkung der Maßnahmen zu erreichen oder zu spät, um die Politik zu beeinflussen
- Aus Bewertungssicht – je länger die Zeit die vergangen ist, je schwieriger ist es die ‚wirklichen‘ Effekte von anderen Faktoren zu isolieren - ausreichende Zeit jedoch erforderlich damit Ergebnisse sichtbar sind
- Viele überprüfte Evaluierungen bewerten geschätzte Auswirkungen eher als tatsächliche, da letztere noch eintreten müssen – begrenzte Nützlichkeit, außer im *Ex-ante*-Sinn
- Um dies zu überwinden, könnten Evaluierungen von Maßnahme in vergangener Programmperiode durchgeführt werden, wenn sich die Maßnahme nicht viel verändert hat; könnte mehr genutzt werden - ist selten
- Auch können erste Glieder der theoriebasierten Kausalkette getestet und geprüft werden, ob vorhergesagte Ergebnisse in der Praxis offensichtlich sind (testen von ‚Vertrauenswürdigkeit‘)

Einsichten - Bedeutung der vorausschauenden Planung

- Eine wichtige Einsicht, die aus den Helpdesk-Aktivitäten gezogen werden muss, ist die Bedeutung der vorausschauenden Planung – zur Ermittlung der durchzuführenden Evaluierungen, ihres Zeitplans, der zu stellenden Hauptfragen, der anzuwendenden Methoden und der zu verwendenden Daten
- Bewertungspläne ermutigten zukunftsorientiertes Denken – Evaluierungspläne von 2014-2020 Programmen förderte dies
- Viele VB versäumten jedoch, erforderlichen Daten so detailliert zu identifizieren, dass diese während der Programmumsetzung hätten erfasst werden können
- Idealerweise sollten Evaluierungsaktivitäten und Datenbedarf als integraler Bestandteil der Programmgestaltung geplant werden
- Um dies zu erreichen, muss die Bewertung jedoch von den politischen Entscheidungsträgern ernst genommen werden
- Dies wiederum erfordert, dass Evaluierungen Ergebnisse liefern, die für die Verbesserung der Politik als nützlich angesehen werden – was wiederum erfordert, dass die am besten geeigneten Methoden und relevantesten Daten verwendet werden. Das bedeutet vorausschauende Planung

**Vielen Dank
für Ihre Aufmerksamkeit**