

fteval JOURNAL

for Research and Technology Policy Evaluation

ISSUE 41 | MARCH 2016

KATHARINA WARTA & STEFAN
PHILIPP

WAS BRINGT DIE
PLATTFORM FTEVAL
UND ZU WAS HAT SIE
ES GEBRACHT?
RÜCKBLICK UND
ERGEBNISSE EINER
MITGLIEDERBEFRAGUNG

MARGHERITA RUSSO, ANNALISA
CALOFFI, FEDERICA ROSSI,
VALENTINA FIORELMONDO &
STEFANO GHINOI

EVALUATING THE
PERFORMANCE OF
INNOVATION
INTERMEDIARIES:
INSIGHTS FROM THE
EXPERIENCE OF TUSCANY'S
INNOVATION POLES

MICHAEL ROTHGANG &
BERNHARD LAGEMAN

EVALUATING CLUSTER
INITIATIVES:
AGENCY,
ORGANISATION,
FUNCTIONALITY,
PERFORMANCE

FEDERAL MINISTRY FOR SCIENCE,
RESEARCH AND ECONOMY

Minoritenplatz 5, 1014 Vienna

Mag.^a Irene Danler

E: irene.danler@bmwfw.gv.at

Mag.^a Simone Mesner

E: simone.mesner@bmwfw.gv.at

Stubenring 1, A -1014 Wien

Mag.^a Sabine Pohoryles-Drexel

E: sabine.pohoryles-drexel@bmwfw.gv.at

FEDERAL MINISTRY OF TRANSPORT,
INNOVATION AND TECHNOLOGY

Radetzkystraße 2, 1030 Vienna

Dr. Rupert Pichler

E: rupert.pichler@bmvit.gv.at

Dr. Mario Steyer

E: mario.steyer@bmvit.gv.at

AQ AUSTRIA

Renngasse 5, 1010 Vienna

Dr.ⁱⁿ Elisabeth Froschauer-Neuhauser

E: elisabeth.froschauer-neuhauser@aq.ac.at

Dr.ⁱⁿ Eva Maria Freiberger

E: eva.maria.freiberger@aq.ac.at

AIT - AUSTRIAN INSTITUTE OF
TECHNOLOGY

Tech Gate Vienna,
Donau-City-Straße 1, 1220 Vienna

Mag. Michael Dinges

E: michael.dinges@ait.ac.at

Mag.^a Barbara Heller-Schuh

E: barbara.heller-schuh@ait.ac.at

AWS - AUSTRIA WIRTSCHAFTSSERVICE
GESELLSCHAFT MBH

Walcherstraße 11A, 1020 Vienna

Dr. Joachim Seipelt

E: j.seipelt@awsq.at

Mag. Norbert Knoll

E: n.knoll@awsq.at

CDG - CHRISTIAN DOPPLER RESEARCH
ASSOCIATION

Boltzmanngasse 20, 1090 Vienna

DIⁱⁿ Brigitte Müller

E: mueller@cdg.ac.at

CONVELOP COOPERATIVE KNOWLEDGE
DESIGN GMBH

Bürgergasse 8-10/I, 8010 Graz

Mag. Markus Gruber

E: markus.gruber@convelop.at

FFG - AUSTRIAN RESEARCH
PROMOTION AGENCY

Sensengasse 1, 1090 Vienna

DIⁱⁿ Dr.ⁱⁿ Sabine Mayer

E: sabine.mayer@ffg.at

Mag. Leonhard Jörg

E: leonhard.joerg@ffg.at

FWF - AUSTRIAN SCIENCE FUND

Sensengasse 1, 1090 Vienna

Dr. Falk Reckling

E: falk.reckling@fwf.ac.at

Dr.ⁱⁿ Dorothea Sturn

E: dorothea.sturn@fwf.ac.at

IHS - INSTITUTE FOR ADVANCED
STUDIES

Josefstädter Straße 39, 1080 Vienna

Mag. Richard Sellner

E: richard.sellner@ihs.ac.at

Dr.ⁱⁿ Brigitte Ecker

E: brigitte.ecker@ihs.ac.at

JOANNEUM RESEARCH
FORSCHUNGSGESELLSCHAFT MBH

Sensengasse 1, 1090 Vienna

Mag. Wolfgang Polt

E: wolfgang.polt@joanneum.at

Mag. Jürgen Streicher

E: juergen.streicher@joanneum.at

AUSTRIAN INSTITUTE FOR SME
RESEARCH

Gusshausstrasse 8, 1040 Vienna

Dr.ⁱⁿ Sonja Sheikh

E: s.sheikh@kmuforschung.ac.at

LUDWIG BOLTZMANN GESELLSCHAFT

Nußdorfer Str. 64, 1090 Vienna

Dr. Peter Mayrhofer

E: peter.mayrhofer@lbg.ac.at

AUSTRIAN COUNCIL FOR RESEARCH
AND TECHNOLOGY DEVELOPMENT

Pestalozzigasse 4/DG 1,
1010 Vienna

Dr. Johannes Gadner

E: j.gadner@rat-fte.at

Dr.ⁱⁿ Constanze Stockhammer

E: c.stockhammer@rat-fte.at

ÖAW - AUSTRIAN ACADEMY OF
SCIENCE

Dr. Ignaz Seipel-Platz 2, 1010 Vienna

Nikolaus Göth, MSc

E: nikolaus.goeth@oeaw.ac.at

TECHNOPOLIS
AUSTRIA

Rudolfsplatz 12/11, 1010 Vienna

Mag.^a Katharina Warta

E: warta@technopolis-group.com

WIFO - AUSTRIAN INSTITUTE OF
ECONOMIC RESEARCH

Arsenal, Objekt 20, PF 91, 1103 Vienna

Dr. Jürgen Janger

E: Juergen.Janger@wifo.ac.at

Dr. Andreas Reinstaller

E: andreas.reinstaller@wifo.ac.at

WWTF - VIENNA SCIENCE AND
TECHNOLOGY FUND

Schlickgasse 3/12, 1090 Vienna

Dr. Michael Stampfer

E: michael.stampfer@wwtf.at

Dr. Michael Strassnig

E: michael.strassnig@wwtf.at

WIRTSCHAFTSAGENTUR WIEN.
EIN FONDS DER STADT WIEN

Mariahilferstraße 20, 1070 Vienna

Robert Mayer-Unterholzer

E: mayer-unterholzner@wirtschaftsagentur.at

ZSI -

CENTRE FOR SOCIAL INNOVATION

Linke Wienzeile 246, 1150 Vienna

Dr. Klaus Schuch

E: schuch@zsi.at

Mag.^a Elke Dall

E: dall@zsi.at

INHALT

ISSUE 41 | 2016

- 3 VORWORT
KLAUS SCHUCH UND STEFAN PHILIPP
- 5 WAS BRINGT DIE PLATTFORM FTEVAL UND
ZU WAS HAT SIE ES GEBRACHT?
RÜCKBLICK UND ERGEBNISSE EINER MITGLIEDERBEFRAGUNG
KATHARINA WARTA UND STEFAN PHILIPP
- 15 EVALUATING THE PERFORMANCE OF INNOVATION
INTERMEDIARIES: INSIGHTS FROM THE EXPERIENCE OF
TUSCANY'S INNOVATION POLES
MARGHERITA RUSSO, ANNALISA CALOFFI, FEDERICA ROSSI, VALENTINA FIORELMONDO
AND STEFANO GHINOI
- 25 EVALUATING CLUSTER INITIATIVES: AGENCY, ORGANISATION,
FUNCTIONALITY, PERFORMANCE
MICHAEL ROTHGANG AND BERNHARD LAGEMAN
- 35 AGENCIFICATION REVISITED
A CASE FOR THE EVALUATION OF FFG AS AN AGENCY
RUPERT PICHLER
- 40 A REGIONAL FUNDING ORGANIZATION FOR BASIC RESEARCH?
UNCONVENTIONAL MISSIONS AND THEIR EVALUATION
GRIT LAUDEL AND MICHAEL STRASSNIG
- 48 ZWISCHENEVALUIERUNG DER VOM BMWF BEAUFTRAGTEN
REGIONALEN KONTAKTSTELLEN – ODER DIE EVALUIERUNG
VON INTERMEDIÄREN STRUKTUREN IN IHREM KONTEXT
BARBARA GOOD
- 53 EVALUIERUNG VON „YOUNG SCIENCE“ – EINEM PROJEKT AN
DER SCHNITTSTELLE ZWISCHEN SCHULE UND WISSENSCHAFT
CAROLINE MANAHL UND SUSANNE DOBNER

- 58 WHEN TWO INTERMEDIARIES AND ONE FUNDING
MINISTRY MEET FOR A COMMON GOAL: THE CASE OF
THE FEDERAL AUSTRIAN INITIATIVE "EVOLVE" FOR
THE CREATIVE INDUSTRIES
ALFRED RADAUER AND TOBIAS DUDENBOSTEL
- 63 AUSWAHLVERFAHREN VON
FORSCHUNGSFÖRDERUNGSORGANISATIONEN IM
INTERNATIONALEN VERGLEICH
PETER BIEGELBAUER UND THOMAS PALFINGER
- 69 „EVALUATION VS. EDUCATED GUESS“
PROJEKTTRÄGER ALS UNTERSCHÄTZE QUELLE FÜR
DIE BEWERTUNG VON FÖRDERPROGRAMMEN
JAN WESSELS, CHRISTIANE KERLEN UND SONJA KIND

VORWORT

KLAUS SCHUCH, Geschäftsführer der fteval und STEFAN PHILIPP, Assistenz der Geschäftsführung der fteval

LIEBE LESERIN, LIEBER LESER,

Das aktuelle Heft legt den Fokus auf die Arbeit von intermediären Organisationen innerhalb eines Forschungs- und Innovationssystems. Es gibt verschiedene Kategorien um Intermediäre zu systematisieren. Als einfacher Einstieg erscheint eine Negativabgrenzung, die auf der Annahme beruht, dass Intermediäre zwischen zumindest zwei, wenn nicht sogar mehreren Systemelementen medieren, also als Vermittler auftreten. Im Kontext der Triple- und Quadrupel-Helix-Modelle, werden intermediäre Organisationen als Agenten beschrieben, deren Tätigkeitsfeld in den Spannungsfeldern zwischen den Systemelementen Wirtschaft, Wissenschaft, Verwaltung und Gesellschaft liegt. Da diese Organisationen einen Beitrag zur Interaktion dieser Akteursgruppen einnehmen und dadurch die Risiken eines Markt- oder Systemversagens mitgieren, werden sie in die Analyse von Forschungs- und Innovationssystemen miteinbezogen. Die vorliegende Ausgabe des fteval Journals for Research and Technology Policy Evaluation stellt die Evaluation der Tätigkeiten dieser Akteure in den Mittelpunkt.

Im ersten Beitrag geht es um eine im internationalen Vergleich eher untypische intermediäre Organisation, die nicht zwischen den großen Systemelementen vermittelt, sondern zwischen Auftraggebern von Evaluationen, Evaluationsdienstleistern und von Evaluation betroffenen F&E-Einrichtungen. Unter dem Titel „Was bringt die Plattform fteval und wozu hat sie es gebracht? - Rückblick und Ergebnisse einer Mitgliederbefragung“ untersuchen Katharina Warta und Stefan Philipp die Rolle der Plattform in der Entwicklung einer Evaluierungskultur in der österreichischen FTI-Politik. Basierend auf einer Rückschau auf die Entwicklungen der letzten 20 Jahre, werden die Herausforderungen vor denen wir in diesem Politikfeld stehen herausgearbeitet.

In den beiden nächsten Beiträgen steht die Vermittlung von F&E und Innovationsüberleitung zur Wirtschaft im Mittelpunkt. Im Beitrag über „Evaluating the performance of innovation intermediaries: insights from the experience of Tuscany's innovation poles“ beschreiben Margherita Russo et al. die Erfahrungen, die im Zuge der Evaluierung von Intermediären und der zugrundeliegenden Indikatorik, gewonnen wurden. Michael Rothgang und Bernhard Lageman thematisieren in Ihrem Beitrag „Evaluating Cluster Initiatives: Agency, Organisation, Functionality, Performance“ die wichtigsten Bausteine einer kritischen Analyse von Clusterpolitiken.

Die jeweils spezifischen Aufgaben und Rollen von Forschungsförderungsorganisationen werden in den Beiträgen von Rupert Pichler beziehungsweise Grit Laudel und Michael Strassnig thematisiert. Kon-

kret hinterfragt Rupert Pichler in seinem Beitrag über „Agencification revisited - A case for the evaluation of FFG as an agency“, ob die Ziele des New Public Management, die der Gründung der FFG zugrunde liegen, erreicht wurden und wirft Fragen auf, die im Rahmen einer Evaluierung der österreichischen Agenturen gestellt werden könnten. In Ihrem Beitrag „A regional funding organization for basic research? Unconventional missions and their evaluation“ beschreiben Grit Laudel und Michael Strassnig die Ergebnisse der Evaluierung des Wiener Wissenschaft- und Technologiefond und die Rolle, die der Fond für die Grundlagenforschung im Wiener FTI-System spielt.

Die nächsten drei Beiträge haben ebenfalls einen spezifischen Bezug zu Österreich. Sie beleuchten erstens die Arbeit von Intermediären an der Schnittstelle zwischen europäischer Forschungsförderung und regionaler Vermittlung, zweitens an der Schnittstelle zwischen Wissenschaft und Gesellschaft und, drittens, das Zusammenwirken von mehreren Intermediären in Bezug auf einen schwer greifbaren und noch vergleichsweise wenig institutionalisierten Bereich, der seinem Selbstverständnis nach an der Schnittstelle zwischen Kunst und Wirtschaft liegt.

Konkret beschreibt Barbara Good in Ihrem Artikel „Zwischenevaluierung der vom BMWF beauftragten Regionalen Kontaktstellen – oder die Evaluierung von intermediären Strukturen in ihrem Kontext“ den Beitrag, den regionale Beratungsservices zur Optimierung der österreichischen Beteiligung an den EU-Forschungsprogrammen spielen. Caroline Manahl und Susanne Dobner berichten über die Ergebnisse der Evaluierung von „Young Science“ – einem Projekt, das an der Grenze zwischen zwei verschiedenen Bildungssystemen (Schule, Wissenschaft) implementiert wurde. Im Beitrag über die Evaluierung der Kreativwirtschaftsinitiative evolve von Alfred Radauer und Tobias Dudenbostel „When two intermediaries and one funding ministry meet for a common goal: The case of the Federal Austrian Initiative “evolve” for the Creative Industries“ wird das Zusammenspiel verschiedener Akteure zur Unterstützung der Kreativwirtschaft beleuchtet.

Die beiden letzten Artikel thematisieren maßgebliche Herausforderungen von Forschungsförderungsorganisationen, nämlich bestmögliche Verfahren zur Bewertung von Projekten und Förderprogrammen. Peter Biegelbauer und Thomas Palfinger fassen in Ihrem Beitrag „Auswahlverfahren von Forschungsförderungsorganisationen im internationalen Vergleich“, die Ergebnisse einer Studie, die diese Auswahlverfahren in Ländern, die zu den EU Innovation Leadern und Followern gehören, zusammen. Im Beitrag von Jan Wessels, Christiane Kerlen

und Sonja Kind über „*Evaluation vs. educated guess - Projektträger als unterschätzte Quelle für die Bewertung von Förderprogrammen*“ wird der Frage nachgegangen, an welchen Stellen extern generierte Evaluationsergebnisse zusätzlichen Erkenntnisgewinn für die Programmverantwortlichen bieten.

Unser herzlicher Dank geht wie immer an die AutorInnen, an die ReviewerInnen für Ihren unermüdlichen Einsatz zur Sicherung der Qualität, sowie an die Graphikerin Caroline Muhr.

Wir wünschen Ihnen viel Vergnügen, Freude und Erkenntnisgewinn bei der Lektüre der Beiträge. Weiters legen wir Ihnen einen Besuch der internationalen Konferenz „*Open Evaluation 2016*“, die am 24. und 25. November 2016 in Wien stattfinden wird, ans Herz. Diesen spannenden Event sollten Sie nicht versäumen; wir freuen uns Sie dort begrüßen zu dürfen!

Ihre

Klaus Schuch

Stefan Philipp

WAS BRINGT DIE PLATTFORM FTEVAL UND ZU WAS HAT SIE ES GEBRACHT? RÜCKBLICK UND ERGEBNISSE EINER MITGLIEDERBEFRAGUNG

KATHARINA WARTA UND STEFAN PHILIPP

1. HINTERGRUND UND VORGANGSWEISE DES INTERNEN ASSESSMENTS DER PLATTFORM

Die fteval wird 2016 zwanzig Jahre alt: Wir nehmen dies zum Anlass, Bilanz zu ziehen und gemeinsam mit den Mitgliedsorganisationen zu erkunden, wodurch und inwiefern die Plattform ihren Mitglieder von Nutzen ist. In der Gründungsphase der fteval gegen Ende der 90er Jahre galt es vorrangig in Österreich eine Evaluierungskultur aufzubauen. Mittlerweile gibt es diese Evaluierungskultur, daher verschiebt sich die Aufgabenstellung der Plattform dahin, diese Kultur weiter zu entwickeln, zu erweitern, zu schärfen, aber auch zu hinterfragen. Mit der Reifung der Evaluierungskultur und den mittlerweile etablierten Kompetenzen bei allen Mitgliedsorganisationen und neuen Rahmenbedingung verändern sich auch die Ansprüche an Evaluierungen. Die Plattform ist ein Ort, um diese Veränderungen zu beobachten und zu begleiten. Wir können davon ausgehen, dass es einen breiten Grundkonsens über die Sinnhaftigkeit der Plattform gibt, da ihre Tätigkeit regelmäßig auf breite Zustimmung stößt. Mit diesem Projekt wollen wir unser Verständnis über die Rolle und den Nutzen der Plattform vertiefen, um bei Bedarf das Angebot, aber auch die Nachfrage entsprechend abzustimmen.

Es gilt zu unterscheiden, was von den Mitgliedsorganisationen allgemein gutgeheißen wird, und was konkret bei der eigenen Evaluierungspraxis als Auftraggeber/in, Evaluator/in oder Evaluierter hilfreich war und ist.

Hierfür sind wir in dreierlei Weise vorgegangen:

- In einem ersten Schritt bzw. in der Geschichte von fteval erstmals überhaupt haben wir den gesamten **Leistungskatalog der Plattform** seit ihrer Gründung erstellt. Die Ergebnisse sind in Abschnitt 3 zusammengefasst und näher analysiert.
- Der zweite Schritt umfasst eine **Mitgliederbefragung**: In zwölf Mitgliedsorganisationen wurden im Herbst 1,5-2-stündige Kleingruppeninterviews mit zwei bis drei Vertreter/innen der Organisation geführt. Die Gruppen setzten sich aus dem/r Vertreter/in

der Organisation in der Plattform (Generalversammlung) und ein bis zwei weiteren Personen zusammen, die in Evaluierungsfragen aus unterschiedlicher Perspektive eingebunden sind, z.B. im internen Monitoring oder als Programmverantwortliche. Alle befragten Personen zählen damit auch zum (potentiellen) Kunden- bzw. Teilnehmerkreis von Veranstaltungen, Publikationen und anderen Leistungen der fteval, wobei zumindest eine der drei Personen nicht unmittelbare Kontaktperson der Plattform ist. Ziel der Kleingruppeninterviews war es, die Erfahrung der Institution zu erfassen, also über die Erfahrung von einzelnen Personen hinauszugehen. Die Fragen umfassten den Bezug der Organisation zu Evaluation im Allgemeinen, die Wahrnehmung und Einschätzung aktueller Entwicklungen und Herausforderungen im Bereich der FTI-Politikevaluierung, die Erfahrung mit der fteval, die Nutzung ihrer Angebote, den Mehrwert dieser Angebote und schließlich auch die Anliegen an die Plattform. Die Ergebnisse dieser Interviews werden in Abschnitt 4 dargestellt.

- Schließlich nahmen wir im November 2015 an der internationalen Konferenz der American Evaluation Association (AEA) teil, stellten dort die Plattform vor und machten bei dieser Gelegenheit auch auf unsere geplante Konferenz zum Thema „Open Evaluation“ in Wien im November 2016 aufmerksam. Rückmeldungen unserer Gesprächspartner fließen ebenfalls in die Analyse der Wahrnehmung der Plattform ein.

2. FTEVAL: EINE EVALUIERUNGS-PLATTFORM MIT EINEM ENGEN INHALTLICHEN FOKUS UND EINEM BREITEN SPEKTRUM AN MITGLIEDERN

Die Tätigkeit der Plattform begann 1996 mit einem Newsletter¹ zu ihrer Gründung als „lose Abfolge von Veranstaltungen / Workshops sowie einer

informellen Publikationsreihe zu Fragen der Evaluierung v. a. von Technologieprogrammen, -projekten und -konzepten.“

Dabei geht es „um das Erreichen eines besseren gemeinsamen Verständnisses zwischen Planern, Förderern und Evaluatoren.“ Damit verbunden waren unterschiedliche Tätigkeiten, nämlich (i) die Diskussion und Entwicklung von best practices und die Diskussion von Methodenfragen, (ii) die Verbreitung von Expertenwissen aus anderen europäischen Ländern und ein Dialog über unsere Grenzen hinweg, (iii) und die Überprüfung und Diskussion der Geltung / Bedeutung von Standards der Europäischen Union für Österreich. In der zweiten Ausgabe des Newsletters (1997) formulieren dieselben Autoren/innen die veränderten Rahmenbedingungen, die sie zur Gründung der fteval motiviert hatten: „FTE-Evaluierungen als Projekt-, Programm- und Institutionenbewertung fristen in Österreich ein vergleichsweise trauriges Dasein. Das gilt für alle Phasen und Formen von Evaluierungen (Ex-ante, Ex-post, Monitoring) ebenso wie für die Auswirkungen, die getätigte Bewertungen auf den Politikformulierungsprozeß haben. (...) Dieses sehr wohl nachdenklich stimmende Bild kommt zumindest von den Voraussetzungen und Rahmenbedingungen her ins Rutschen: Vor allem kommen seit dem EWR/EU-Beitritt neue Qualitätsstandards auch in den Bereich der FTE-Evaluierungen. (...) Auf nationaler Ebene zwingen die knapper fließenden Ermessensausgaben (...) und der stärker spürbare Verteilungswettkampf zu einem bewußter geplanten Umgang mit Geld.“ Es folgt eine Reihe von Zielen für die Plattform, die die Er-

höhung der Verbindlichkeit und der Regelmäßigkeit von Evaluierungen, die laufende Diskussion, Anpassung und Optimierung der Methoden, die Verbesserung des Zusammenhangs von Programm- bzw. Institutionendesign und Evaluierung, Offenheit und Blick von außen, Verbesserung der in Österreich sehr präkeren Indikatorensituation als eine wichtige Grundlage für Evaluierungen, Vorantreiben der Diskussion über einen 'code of conduct' bezüglich Datenschutz und für die Unionsebene den Aufbau von Expertise bei Beamten und Evaluierungsspezialisten umfasst.

Heute sind die Ziele der Plattform wie folgt definiert²: „Mission der Österreichischen Plattform für Forschungs- und Technologiepolitikevaluierung ist es, mehr, bessere, transparentere Evaluierungen für eine bessere strategische Planung der FTE-Politik in Österreich zu erreichen – also gemeinsam mit den österreichischen technologie- und forschungspolitischen Entscheidungsträgern Evaluierungskultur zu entwickeln.“

Der Fokus hat sich – was auch im Namen ersichtlich ist – auf Forschungspolitik erweitert und insgesamt spricht man mehr von Innovations- als von Technologiepolitik. Die Aufgaben der Plattform jedoch sind im Wesentlichen dieselben geblieben.

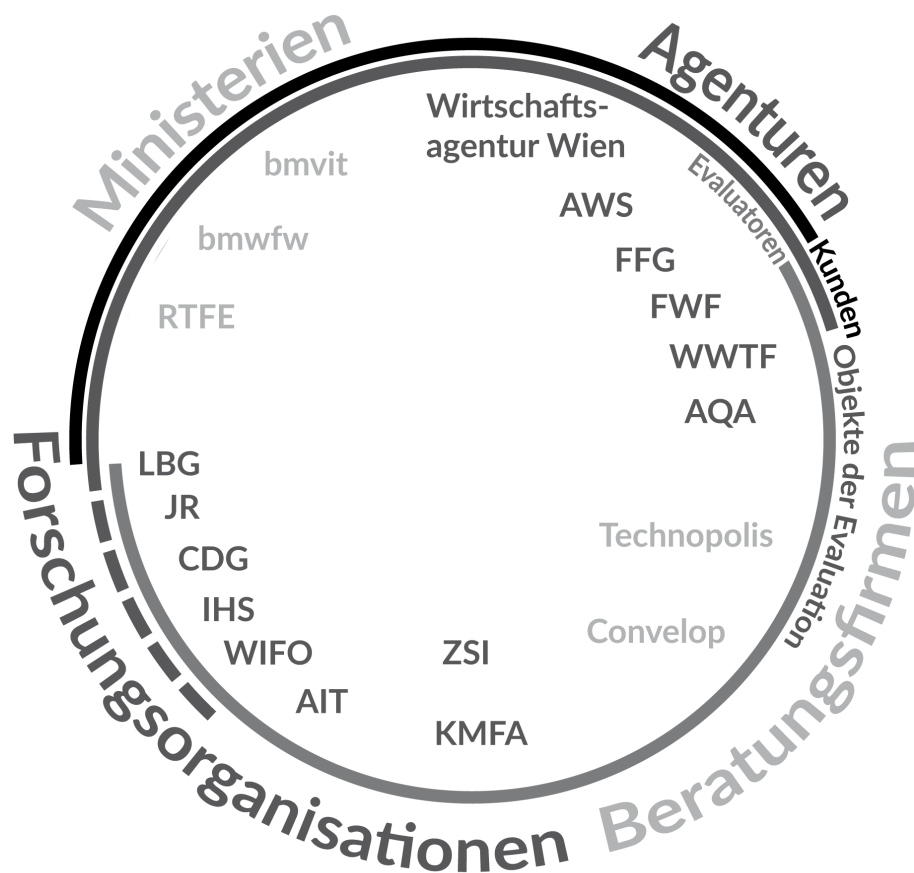


Abbildung 1: Mitglieder der fteval: Stakeholdergruppen und ihre Rollen in Bezug auf Evaluierung

Betreffend der Forschungsorganisationen gibt es zwei Aspekte: Sie sind Evaluatoren (AIT, WIFO, ZSI, JR, KMFA und IHS) oder sie evaluieren in eigener Sache (LBG und CDG). In diesen Funktionen sind sie auch in der fteval vertreten, also jeweils mit ihrer strategischen Abteilung bzw. ihrer Expertise als Evaluatoren/innen, nicht jedoch als evaluierte oder zu evaluierende Organisationen. Hier spiegelt sich sowohl die ursprüngliche Ausrichtung der Plattform wieder – bei der zwei der Forschungsorganisationen und ein Ministerium federführend waren – als auch die historisch unterschiedliche Entwicklung der Finanzierung und Governance institutioneller Forschung und projektbasierter Förderung.

3. IM 20. JAHR: RÜCKSCHAU AUF DIE ENTWICKLUNG DER ANGEBOTE

Die fteval hat ein ausdifferenziertes Set an Leistungen entwickelt, das im Zuge dieser Untersuchung erstmals zur Gänze erhoben und ausgewertet wurde. Die Ergebnisse sind in der folgenden Abbildung zusammengefasst und in der Folge im Detail erläutert. Die Abbildung stellt die Entwicklung der fteval über die Zeit dar, und bildet im Säulendiagramm die Zahl an Mitgliedern, unterschieden nach dem Typus der Organisation, dar. Darunter sind ausgewählte Publikationen (das fteval Journal, früher Newsletter genannt) und Veranstaltungen pro Jahr gezählt, weiters werden gewichtigere Events gesondert ausgewiesen, nämlich Konferenzen, Trainings und die Publikation der Standards. Schon auf den ersten Blick wird deutlich, dass sich die fteval hinsichtlich der Mitglieder über zehn Jahre in einer Expansionsphase auszeichnete und seit 2010 konsolidiert hat. Weiters sieht man anhand der Darstellung die hohe Kontinuität der Aktivitäten, die für solche Vereine nicht selbstverständlich ist.

Der erste Newsletter im Jahr 1996 wurde noch von Vertreter/innen dreier Organisationen (dem jetzigen BMVIT, Joanneum Research und WIFO) herausgegeben, der schon nach wenigen Jahren war die Zahl der Mitglieder doppelt und dreifach so groß. So waren bei der Institutionalisierung der bis dahin losen Kooperation durch die Gründung einer Gesellschaft nach bürgerlichem Recht im Jahr 2002 bereits zwölf Organisationen Mitglied; bei Gründung des Vereins im Jahr 2006 waren es bereits 16. Im Jahr 2015 hat die Plattform 19 Mitglieder und durch den beschlossenen Beitritt der Österreichischen Akademie der Wissenschaften im Jahr 2016 steigt die Mitgliederzahl auf zwanzig³.

Einhergehend mit der Institutionalisierung und der Ausweitung der Mitgliederbasis stieg auch das verfügbare Budget der Plattform zur Durchführung ihrer Aktivitäten an. Während der vergangenen Jahre lag das Budget zwischen einem Minimalwert von rund 45.000 Euro (2011) und einem Ausreißer im Jahre 2006 von rund 170.000 Euro (das Jahr des österreichischen EU-Ratsvorsitzes und damit in Verbindung die international ausgerichtete Konferenz zu „New Frontiers in Evaluation“ in Wien). Diese Finanzmittel sichern den Betrieb der Geschäftsstelle, die Publikation des fteval Journals, die Organisation und Durchführung von verschiedenen thematischen Veranstaltungen und anderen Aktivitäten wie z.B. das Anbieten von Trainingsworkshops oder auch die Organisation internationaler Konferenzen.

Die **Herausgabe des fteval Journal**, das aus dem früheren Newsletter hervorgegangen ist, ist Ausgangspunkt und damit ältestes Produkt der Plattform. Seit der Gründung der fteval wurden insgesamt 40 Ausgaben dieser Reihe veröffentlicht, die einerseits Rahmen für wissenschaftlichen Diskurs und andererseits Plattform für praxisnahen Erfahrungsaustausch ist. In diesen Ausgaben wurden insgesamt 184 Artikel publiziert, von denen drei Viertel in englischer Sprache verfasst wurden.

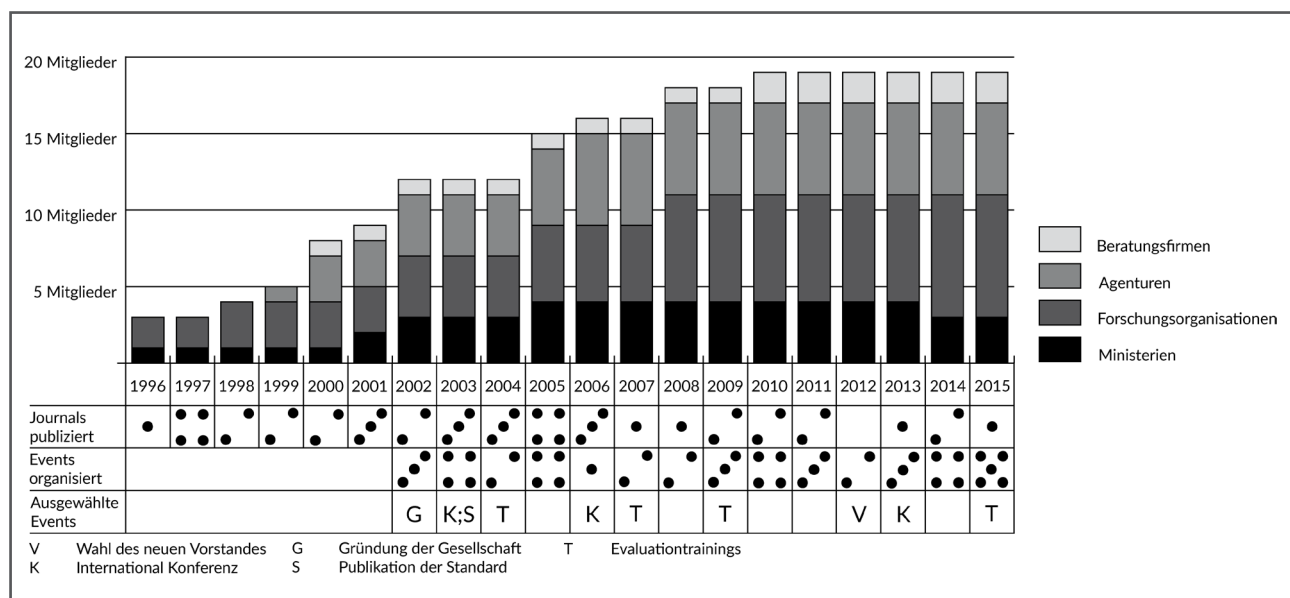


Abbildung 2: Die Entwicklung der fteval: Mitglieder und ausgewählte Aktivitäten

3 Bei der Entwicklung der Mitgliedszahlen ist zu berücksichtigen, dass durch die Fusion von FFF (Forschungsförderungsfonds) und TIG (Technologie-Impulsgesellschaft) im Jahr 2004 und die Zusammenführung von Wissenschafts- und Wirtschaftsministerium zum BMWF (Bundesministerium für Wissenschaft, Forschung und Wirtschaft) im Jahr 2015 Mitglieder „wegfielen“, ohne dass Akteure die Plattform verlassen haben.

Obwohl es sich beim fteval Journal um ein Publikationsorgan handelt, das vor allem von den Mitgliedern getragen wird, ist es immer gelungen, auch Autoren/innen aus dem Ausland für dieses Format zu gewinnen; so stammen rund 40% der Artikel von dieser Autoren/innengruppe.

Einen wichtigen Baustein zur Weiterentwicklung der Evaluationskultur stellen die **fteval Standards** der Evaluierung in der Forschungs- und Technologiepolitik dar. Diese dienen dazu, sowohl den Evaluatoren/innen, den beauftragenden Einrichtungen als auch den Evaluierten einen Referenzrahmen zur Verfügung zu stellen. Damit soll ein höherer Grad von Verbindlichkeit und Sicherheit für alle Beteiligten erreicht werden. Die Standards wurden in einem gemeinsamen Prozess, der mit der Institutionalisierung der fteval im Jahr 2002 startete, von den Mitgliedsorganisationen entwickelt und als gemeinsame Mindestvoraussetzung für Evaluierungen beschlossen. Die aktuelle Version der Standards⁴ (2012) geht ebenfalls aus einem gemeinschaftlichen Prozess der Überarbeitung hervor.

Weitere zentrale Tätigkeiten der Plattform stellen die **Sammlung und Veröffentlichung der österreichischen Evaluationsberichte** aus dem Bereich der FTI-Politik⁵ dar. Durch die Abrufbarkeit von rund 300 Evaluationsberichten, Studien und Strategiedokumenten auf der Website der Plattform wird ein Beitrag zur Transparenz und Sichtbarkeit der Aktivitäten in diesem Politikfeld geleistet.

Das Angebot von **Trainingsworkshops** leistet einen wichtigen Beitrag zur Weiterentwicklung und zum Aufbau von Know-how im Bereich der FTI-Politikevaluierung. Die ersten (dokumentierten) Trainings fanden im Jahr 2004 statt, als Trainer/innen der Plattform direkt an sieben verschiedenen Institutionen ein- bis dreitägige Workshops abhielten. In den Jahren 2007 und 2009 wurden dann Methodenschulungen zu den Themen Bibliometrie, Surveys und Arbeiten mit Daten, Soziale Netzwerkanalyse, Peer Review/ Expert Panel, Focus Groups, ökonomische Methoden der Wirkungsanalyse, makroökonomische Modelle sowie zu Interviews und Fallstudien mit insgesamt mehr als 140 Teilnehmern/innen angeboten. Nach längerer Unterbrechung folgten im Jahr 2015 Trainings, die speziell für jüngere Mitarbeiter/innen konzipiert waren, zu den Themen ex-ante, interim und ex-post Evaluation und die mit mehr als 80 Teilnehmern/innen ebenfalls sehr gut besucht waren.

Die Plattform organisiert darüber hinaus verschiedenste thematische **Veranstaltungen**. Seit dem Jahr 2002 fanden jedes Jahr bis zu fünf verschiedene Events statt. Diese Veranstaltungen dienen einerseits dem Austausch über abgeschlossene Evaluationen (z.B. der österreichischen Systemevaluierung im Jahr 2009) und der Diskussion von evaluationsrelevanten Entwicklungen (z.B. zur Einführung von ORCID im Jahr 2015) wie auch andererseits der Vernetzung der österreichischen Community in sich sowie mit ExpertInnen aus dem Ausland (z.B. Research Evaluation in Austria mit tschechischen ExpertInnen im Jahr 2015). Die Veranstaltungen werden in der Regel in Kooperation mit einem oder zwei institutionellen Partnern durchgeführt. Eine Besonderheit im Bereich der fteval Veranstaltungen stellt das Format „fteval inside : fteval insight“ dar, bei dem Mitgliedsorganisationen zu sich einladen, um ihre Tätigkeiten mit einem selbstgewählten thematischen Focus zu präsentieren

und in informellem Rahmen zu diskutieren. Dies ist das einzige Format der Plattform, zu dem nur Mitarbeiter/innen der Mitgliedsorganisationen eingeladen werden.

Zusätzlich zu diesen kleineren Veranstaltungen organisierte die Plattform bisher drei international ausgerichtete **Konferenzen** zum Thema FTI-Politikevaluierung in Wien. Die erste Konferenz fand im Jahr 2003 zum Thema „Government funded R&D activities, evaluation and lessons learned“ mit rund 190 Teilnehmern/innen aus dem In- und Ausland statt. Die Konferenz „New Frontiers in Evaluation“ im Jahr 2006 lockte 230 BesucherInnen an. Die Konferenz „New horizons / new challenges – Evaluation of STI policies, instruments and organisations“ (2013) besuchten insgesamt 210 Personen. Damit zählt das Format der fteval Konferenz zu den größten europäischen Konferenzen zum Thema FTI-Politikevaluierung. Da das Feedback der Community die letzten Male sehr positiv war, ist angedacht, die Konferenz in einem Dreijahresrhythmus zu organisieren, die nächste Konferenz findet am 24. und 25. November 2016 mit dem Titel „Open Evaluation“ in Wien statt.

Neben diesen umfassenden Angeboten steht die fteval auch für punktuelle **Beratung** zur Verfügung, z.B. wenn es um Fragen zu Formulierungen bei der Ausschreibung von Evaluierungen geht. Sie betreibt eine Homepage und stellt ihren Mitgliedern Materialien für Marketing bereit. Vor allem die Repräsentant/innen der Plattform – der Geschäftsführer bzw. der oder die Obmann/Obfrau – nehmen auch die Vertretung der fteval nach außen, wie z.B. bei Evaluierungstrainings oder einschlägigen Veranstaltungen anderer Organisationen, wahr.

4. ERGEBNISSE DER MITGLIEDERBEFRAGUNG

Im Folgenden fassen wir die Ergebnisse der Mitgliederbefragung entlang von fünf Kernaussagen zusammen. Diese betreffen (a) die Erfolge beim Aufbau einer Evaluierungskultur, (b) Veränderungen in der Rolle der Plattform, (c) neue inhaltliche Herausforderungen für Evaluierungen sowie (d) die steigende Komplexität von FTI-Politik, und schließlich (e) die Positionierung der fteval als Kommunikationsplattform.

A. AUFBAU EINER EVALUIERUNGSKULTUR

Unter den Mitgliedsorganisationen besteht ein breiter Konsens, dass die fteval viel zur Etablierung einer Evaluierungskultur im Bereich Forschungs- und Technologiepolitik in Österreich beigetragen hat. Die Publikation gemeinsamer Standards im Jahr 2003 wird als zentrales Element genannt, ebenso wie Evaluierungstrainings, Diskussionsveranstaltungen, Konferenzen und last but not least die Publikation der vorliegenden Reihe des fteval Journals, die sowohl die forschende Reflexion über Evaluierung als auch die Praxis aus Politik und Verwaltung adressiert.

„Die Plattform hat mehr Rationalität in das System gebracht. (...) Wir bringen mehr Evidence in den Politikzyklus, sie hat das Thema am Leben erhalten. Es war nie die Gefahr, dass die Community in das zurückfällt, was sie vor 1996 war.“

4 Siehe http://fteval.at/upload/fteval_Standards_deutsch.pdf
5 Forschungs- Technologie- und Innovations-Politik.

„Ich halte es für extrem wichtig, dass die Leute, die in einem Fachgebiet tätig sind, sich regelmäßig treffen. Hier hat die fteval schon das wichtigste gemacht, die Intensität zu erhöhen, ist nicht notwendig. Die Vernetzungsaktivitäten sind extrem wichtig, deswegen ist es gut, dass es die Plattform gibt.“

„Die fteval ist ein Zugangspunkt zu dem Thema. Durch die verschiedenen Formate ist es ein Weg zu beobachten, was sich tut, was bewegt, was bewegt wird. Es ist eine gewisse Form von 'enabling environment', ansonsten müsste man sich andere Orte suchen, oder eine Plattform gründen. Man richtet den Blick darauf und hat gesammelte Informationen. (...) Das Angebot der fteval hat insofern eine Bedeutung, als es hilft, alle bei der Stange zu halten.“

„Es ist schon interessant, wenn die Welten informationsmäßig zusammen kommen und nicht nur, wenn etwas kontrolliert wird.“

„Ich bin unheimlich glücklich über die große internationale Konferenz (...) das ist ein Impact der Plattform, ohne die fteval gäbe es die Organisationskapazität nicht, das zu machen.“

„Die Standards stehen immer unter Feuer, aber sie sind wie ein Zahnbürstl (...), man kann sich darauf berufen und die Übung, darüber zu reden, ist eine der wenigen Erziehungsmöglichkeiten, die wir uns gegenseitig in der Plattform angedeihen können.“

„Die Plattform hat einige (internationale) Impulse aufgegriffen und verstärkt, insbesondere den EU-Beitritt, die einer Modernisierung des Politikfeldes dienlich waren.“

„Ich bin ja auch im internationalen Bereich aktiv. Da wird man auch beneidet.“

Betrachtet man die Frage nach einer gemeinsamen Evaluierungskultur aus der Nähe, zeigt sich jedoch, dass trotz der Bemühung um gemeinsame Standards unterschiedliche Evaluierungskulturen nebeneinander existieren. Auch die Einschätzungen darüber, inwiefern die Standards eine reale Auswirkung auf Evaluierungen haben, sind unterschiedlich, ja teils widersprüchlich.

- Die Interviewten sind sich darin einig, dass der Fokus der Plattform auf Programmen und nicht auf Institutionen liegt, und auf Zwischen- und ex-post-Evaluierungen eher als auf ex-ante Evaluierungen. Begleitende Evaluierung hat in den letzten Jahren vermehrt Eingang gefunden. Hinsichtlich der Evaluierungskultur führt das dazu, dass Programmevaluierungen gut etabliert und tatsächlich weitgehend nach gemeinsamen Standards durchgeführt werden. Die in der Gründungsphase der fteval beschriebenen Probleme⁶ sind im Bereich der Programmevaluierungen überwunden. Es liegt in der fteval jedoch wenig Erfahrung mit institutioneller Evaluierung vor, gleichzeitig gibt es diese Erfahrungen, sie sind jedoch wenig in der Plattform verankert bzw. werden sie von ihr bislang wenig aufgegriffen (dies betrifft insbesondere die Evaluierung akademischer Einrichtungen). Hinsichtlich der begleitenden Evaluierungen gibt

es ein breites Spektrum an Einschätzungen, das von großer Begeisterung wegen ihres inhärenten Lernpotentials bis zur Skepsis reicht, die infrage stellt, ob es sich hier aufgrund der sich im Prozess ergebenden Nähe der Evaluatoren zum Evaluierungsgegenstand nicht eher um einen Coaching-Prozess handle. Die Evaluierungskultur ist also im Kernbereich gefestigt, entwickelt sich aber in anderen Bereichen erst nach und nach aus.

„Es hat sich die Plattform extrem gut eingerichtet, wo es ums Evaluieren von Programmen geht und weniger, wo es um Institutionen-Evaluierung geht. Das hat mit der Besteller-Situation zu tun: da gibt es Peergroups... das kommt gar nicht in die Plattform, das wird selbst bestellt. Entweder es findet nicht statt oder es wird von Qualitätsagenturen gemacht. Trotzdem ist das unbefriedigend, die Kurve ist schon sehr flach, was man noch lernen kann.“

„Die Erwartung von Programmleitern an Evaluierung ist sicher eine Bestätigungsfunktion. Das zweite ist eine Lernfunktion, das ist eine Qualität, die vorher nicht da war.“

„Bei Programm- und Institutionenevaluierung: Das ist nicht nur ein Versäumnis der Evaluierungskultur, das sind auch die Förderinstrumente, die nicht so revolutionär sind.“

„Ex-ante Evaluierung wird mit Argumentationshintergrund verwechselt.“

- In den vergangenen zwei Jahrzehnten hat sich die Evaluierungs-Community insgesamt weiterentwickelt, es gibt heute eine durchaus überschaubare Zahl gut etablierter Evaluierungsorganisationen. Diese haben, so die Wahrnehmung der Befragten, alle mehr oder weniger dieselben Qualifikationen, mit anderen Worten, die Community ist gewachsen, jedoch ohne sich nach Spezialisierungen ausdifferenzieren. Man könnte hierfür die Metapher „viele Generalisten aber keine Fachärzte“ heranziehen.
- Betreffend der Standards sagen die einen, diese seien heute perfekt internalisiert und dementsprechend ein code of conduct für alle Akteure, andere sind deutlich kritischer. Sie beobachten regelmäßig, dass die Standards nicht eingehalten oder umgangen werden, und dass es bislang nicht gelungen ist, solchen Problemen, ja sogar Verletzungen auch tatsächlich nachzugehen, was einer effektiven Qualitätskontrolle entspräche.

„Als die Standards gemacht worden sind, war viel davon NICHT Standard. Jetzt werden sie im Regelfall erfüllt. (...) Das war eine extrem hohe Wirksamkeit.“

„Meine Vermutung ist, dass die Standards unbewusst angewandt werden, aber das ist nicht sicher. (...) beim Wiederlesen bleibt wenig übrig, was man brauchen kann.“

„Die Auftraggeber lesen die Standards nicht oder haben sie nicht verinnerlicht.“

„Wenn wir eine Evaluierung anbieten, steht das im Angebot drin, dass wir uns den Standards verpflichtet fühlen.“

„Die Standards können ein gewisser Orientierungsrahmen sein (...), es ist wichtig, dass es das gibt, um das Feld zu bestimmen. Evaluation ist ja keine Profession.“

- Konkret bedeutet dies, dass beispielsweise die Grenzen zwischen „Studien“ und „Evaluierungen“ verwischt sind, und erstere sich nicht an den Standards orientieren, obwohl sie wesentliche Aspekte von Evaluierungen beinhalten. Insbesondere ex-ante „Evaluierungen“ kommen im österreichischen FTI-Bereich unter diesem Titel kaum vor, was aber nicht heißt, dass die Akteure sich nicht über „Roadmaps“, „Bedarfsstudien“, „Sichtungen“, „Konzepte“ etc. informieren, bevor sie Förderaktivitäten in Angriff nehmen. Transparente Vergabemodalitäten, die für Evaluierungsstudien mittlerweile etabliert sind, werden in diesen Fällen häufig ignoriert, Evaluierungskriterien werden nur fallweise angewendet. Aber auch Zwischenevaluierungen werden bisweilen als oder in „Studien“ verpackt, insbesondere wenn es sich nicht um klassische Programme handelt sondern um komplexere Interventionen, wie umfassende Innovationsstrategien, über deren Umsetzungsgrad „berichtet“ wird.
- Schließlich wurde auch die Frage nach Unabhängigkeit und Interdependenz aufgeworfen: Aus der Akteurskonstellation geht hervor, dass immer wieder die Auftraggeber von Evaluierungen auch Gegenstand der Evaluierung sind, indem sie beispielsweise Zuständigkeiten im Design und in Auswahlverfahren der Programme haben, deren Evaluierung sie beauftragen. Das kann Ministerien ebenso betreffen wie Agenturen. Die Mitgliedsorganisationen der fteval sind sich dieser Herausforderung bewusst und haben unterschiedliche Wege ausprobiert, um dem Anspruch nach Unabhängigkeit gerecht zu werden, zum Beispiel indem sie eine/n unabhängige/n Expertin/en, möglicherweise aus dem Ausland, in das Auswahl- und Begleitkomitee der Evaluierung einbeziehen. Aber auch wenn keine institutionellen Verbindungen bestehen gibt es eine Sorge, Evaluatoren/innen könnten gehemmt sein, kritische Ergebnisse zu formulieren, da sie angesichts des kleinen Marktes und der vorherrschenden Konkurrenz um spätere Aufträge fürchten würden. Konkret konnte jedoch keine solche „Korrektur“ oder sogar vorauseilende Zensur genannt werden, wenn auch Erfahrungen existieren, dass in der Endphase von Evaluierungen gemeinsam mit dem Auftraggeber bzw. dem Begleitkomitee die konkrete Formulierung von Schlussfolgerungen diskutiert und gegebenenfalls auch überarbeitet wird. Im Zusammenhang mit „freundlichen“ Schlussfolgerungen wurde von mehreren Interviewpartnern angemerkt, dass Programmevaluierungen tendenziell „gut ausgehen“, wenn sie ein Programm und nur dieses Programm betrachten, weil, wenn „nicht sehr viel schief“ läuft, mit positiver Additionalität der Förderung zu rechnen ist. So haben insbesondere Programmevaluierungen häufig einen bestätigenden Charakter und dienen nur am Rande dem Lernen. Weiter unten werden wir darauf eingehen, dass angesichts steigender Komplexität und knapper Mittel ein zunehmender Bedarf nach kritischer Betra-

chtung erwartet wird, und zwar mit einer besonderen Aufmerksamkeit für das größere Portfolio.

„Es gibt immer noch ein vollständig fehlendes Governance-Modell von Evaluierung: Wer als Evaluierungsherr auftritt, das wird von den Auftraggebern überhaupt nicht betrachtet.“

„Bei uns ist es die Regel, dass Evaluierungen nicht von uns sondern den zuständigen Programmauftraggebern beauftragt werden. (...) Ein bisschen hinkt die Logik, wenn ich Konzept und Design-Fragen stelle, da die Interventionsstrategie ja auch vom Ministerium formuliert wird.“

„Zur Frage: Wer ist der beste Auftraggeber für eine Evaluierung? Wir versuchen es zu umgehen, weil wir die Evaluierung in die Strategieabteilung legen und eine externe Person aus dem Feld der Partnerorganisationen von Science Europe hinzuziehen.“

B. NACH 20 JAHREN IST DIE ROLLE DER PLATTFORM NEU ZU DEFINIEREN

Wie eingangs erwähnt ergab sich der Fokus der fteval aus den politischen Rahmenbedingungen Mitte/Ende der 90er-Jahre, nämlich dem österreichischen EU-Beitritt, der Einführung (teilweise verpflichtender) Evaluierungen von neuen Förderprogrammen und nicht zuletzt aus der zunehmenden Bedeutung von Innovationspolitik. In diesem Kontext ist eine neue Generation von Praktikern/innen in der Forschungs- und Innovationspolitik sowie von Evaluatoren/innen herangewachsen, die sich stark an internationalen Maßstäben und Diskussionen orientierten. Sie gründeten 1996 die fteval mit dem Ziel, eine Evaluationskultur zu etablieren; das hieß damals, mit einer vorherrschenden Unkultur zu brechen. Heute ist dieser klar revolutionäre Ansatz überholt, die Plattform steht nunmehr vor der Aufgabe, das Erreichte zu pflegen. Wie die Ergebnisse dieser Untersuchungen zeigen, gilt es aber gleichzeitig auch, das Erreichte in Frage zu stellen. Manche Interviewpartner stellen infrage, ob das systemisch überhaupt geht.

„Evaluierungskultur ist immer auch ein Teil von Risikovermeidung und nicht von Systemveränderung. Es passiert auf Basis der Evaluierung zu wenig. Man bräuchte weniger Evaluierungen, die mehr verwendet werden.“

„Am Anfang seid Ihr am wirksamsten gewesen, jetzt könnt Ihr der Treiber der Tradition sein. Vorher ward Ihr provokant. Ihr seid stark im Konstruieren von Standards - könnt Ihr auch stark im Dekonstruieren sein, wo man hinterfragt, wo sind die? Oder treibt Ihr weiter im Mainstream?“

„Es gibt einen Strickfehler im System, es dominieren Legitimation und Bestätigung. (...) Es ist durch das System wenig sichergestellt, dass man unbefangen kritisch drauf schaut. (...) Die Evaluierungen liefern sehr viel Bestätigung, stärken die Programme, aber machen den Ausstieg sehr schwer, die Reduktion der Programmvvielfalt. Da wirken Evaluierung und das System, das evaluiert wird, stabilisierend.“

„Die Plattform ist ein bisschen ritualisiert, aber es ist zu wenig anstößige Diskussion mehr da. Das letzte Mal hat die gleich nahezu zu einer Zerreißprobe geführt, aber so etwas braucht das, eine spannende Diskussion.“

„Wollen wir das System aktivieren, stabilisieren oder provozieren, irritieren? Das kann man nicht zur selben Zeit machen.“

„Crowd funding, open innovation: Es gibt viele neue Formate, Wissenschaft zu generieren, die sich innerhalb der Wissenschaft und rundherum etabliert haben und die erfordern neue und andere Formate an Evaluierungen.“

Nach den einzelnen Angeboten der fteval befragt gibt es eine sehr breite Zustimmung, die Tätigkeiten und Instrumente der Plattform werden also nicht infrage gestellt. Die Herausforderung liegt eher darin, die (selbst-)kritische Kompetenz aufrecht zu halten und ernst zu nehmen.

C. DIE GROSSE HERAUSFORDERUNG LIEGT IM NEUEN BEDARF NACH WIRKUNGSANALYSE

In der ersten Publikation der Plattform schrieb Erik Arnold *„Since an Audit Office's core skill is usually accounting, evaluations tend to use cost-benefit or cost-effectiveness approaches. These work adequately with simple policy questions where costs and benefits can easily be measured and modelled. However, science, technology and innovation actions tend to involve significant uncertainties. Their effects are often indirect, and are difficult to measure, so accounting-based approaches are not very useful.“* Das Österreichische Bundesbudget wird jedoch seit 2013 nach Wirkungszielen geführt, durch Ziele, die von der Verwaltung durch quantitative Indikatoren definiert und regelmäßig evaluiert werden müssen. Das Messen von Effekten und Wirkung ist somit zu einer Verpflichtung geworden.

„Früher haben wir in Evaluierungen auch die Grundlagen aufgearbeitet. Jetzt wird nur noch die pragmatische Frage gestellt: Macht das einen Sinn, was wir da tun? Früher hat man einen Mix aus Evaluierung und Studie gemacht, und ein Programm von allen Seiten beleuchtet. Heute ist der Impact deutlich stärker ins Licht gerückt. Er ist schwer fassbar, weil er methodisch so schwer zu fassen ist. (...) Da gibt es Anbieter, die sich rechtfertigen müssen, gegenüber den Organisationsherren.“

„Es tauchen plötzlich so Begriffe auf wie Kosten-Nutzen-Analyse. Entweder man will das, das heißt methodisch viel. Auch der Wirkungsbegriff: Da weiß man nicht, wie der zu interpretieren ist. Wie steht das in Relation zu anderen Fragen, zu den Ressourcen, zu dem, was als Impact messbar sein kann?“

„Die Durchführung von Programmen ist weitgehend professionalisiert und die Prozesse werden wenig angeschaut. Wie werden Stakeholder und Communities erreicht, adressiert? Was passiert in den Prozessen?“

Diese Verpflichtung fließt nach und nach auch in die „Terms of References“ (ToR, Pflichtenhefte) von Evaluierungen, wie sie die Mitglieder der fteval schon vor der Budgetreform kannten, ein.

Aber auch über diese Verpflichtung hinaus gewinnen Wirkungsanalysen in der Evaluierung an Bedeutung. Das hängt mit der Reife des Systems zusammen, in dem sich die gesamte Informationsbasis über die Gebarung von Programmen durch professionelles (internes) Monitoring verbessert hat. Elektronisches Bewerbungs- und Berichtswesen ermöglicht den Agenturen fast in *real time* über Input- und Output der Programme zu berichten. Längerfristige Wirkungen hingegen lassen sich so nicht erfassen, dies wird also an Evaluationsstudien ausgelagert.

Die Herausforderung für externe Evaluierung wird dadurch größer und komplexer (siehe hierzu auch Punkt d unten). Mit der Frage nach der Wirkungsanalyse geht auch die Wahl des geeigneten Zeitpunkts der Evaluierung einher: Der Politikzyklus verlangt nach frühzeitiger Rückmeldung, sowohl wenn es um eine Art von Bestätigung, als auch wenn es um Kritik und Lernoptionen geht. Zu diesem Zeitpunkt ist es aber für eine Wirkungsanalyse noch zu früh, weil sich erstens die Wirkungen⁷ häufig noch nicht eingestellt haben und zweitens, selbst wenn, Daten hierüber noch nicht zur Verfügung stehen.

D. DIE EVALUIERUNG VON FTI-POLITIK WIRD KOMPLEXER

Der Bedarf, Wirkungen öffentlicher Intervention und Förderung zu verstehen, führt unmittelbar zu einer weiteren Problematik, die von vielen Gesprächspartnern hervorgehoben wurde, nämlich der steigenden Komplexität, in der sich die Akteure im FTI-Bereich bewegen. Das betrifft das Design von Maßnahmen ebenso wie ihre Umsetzung. In Zeiten knapper öffentlicher Budgets wird der Ruf lauter, Interdependenzen zwischen unterschiedlichen Politikbereichen zu berücksichtigen. Trotz der verbesserten Datenlage über Forschung, Innovation und diesbezügliche Förderungen bleibt es schwierig, Ursache-Wirkungsbeziehungen zu verstehen und – wenn möglich quantitativ – zu beschreiben. Wir sind, könnte man mit einem gewissen Augenzwinkern sagen, nach wie vor verwirrt, wenngleich heute auf einem höheren Niveau als früher. Das hat jedoch konkrete Auswirkungen auf Evaluierung:

- Die *Terms of References* sind häufig überladen. Sie werden oft kumulativ geschrieben, d. h. sie enthalten alle Fragen, die bisher in Evaluierungen gestellt wurden, und ein paar neue, für die betroffene Evaluation spezifische Fragen.

„Es kommt nicht nur auf die Qualität von Programmen und Evaluierungen an, sondern auch und vor allem auf das Selbstverständnis auf Auftraggeberseite.“

„Die nachfragenden Akteure sind schlanker und hungriger geworden. Die Anbietenden sind auch nicht fatter geworden und müssen mit mehr Anbietern im eigenen Dunstkreis umgehen.“

„Ich mache eine Evaluierung und schmeiße alles hinein, was ich bisher gefragt habe und gehe viel zu wenig auf die Spezifika ein. Wenn ich in die Auswahl der Evaluatoren eingebunden war, haben mir die Anbieter leid getan, weil die Terms viel zu umfangreich waren und trotzdem wurden manchmal wichtige Dinge ausgelassen. (...) Die Evaluationskultur ist eine, die nach der Vollständigkeit der Evaluierung strebt und weniger nach dem Sinn der Intervention fragt.“

„Die Frage, ließe sich das Problem auch mit anderen Instrumenten abdecken, wird nicht gestellt.“

- Qualitative Methoden stoßen auf zunehmendes Interesse und werden als profunde analytische Ergänzung zu quantitativen Methoden genutzt, nicht nur um Ergebnisse zu illustrieren. Damit steigen auch die Ansprüche an die Professionalität qualitativer Forschung.

- Auch die quantitative Analyse von Förder- und Wirkungsdaten verändert sich. Früher waren Evaluierungen die erste Gelegenheit, um einen Überblick über Fördertätigkeit und Zielgruppen sowie deren Bewertung der Förderung zu erstellen, letztere wurden mittels Fragebögen erhoben. Heute kennen die Agenturen ihre Förderdaten besser (früher gab es beispielsweise kaum Zeitreihen, heute ist das eine Selbstverständlichkeit), ebenso wie die zuständigen Ministerien, die regelmäßige Updates einholen. Von Evaluatoren erwartet man sich heute nicht nur, dass sie Zielgruppen über standardisierte Erhebungen befragen, sie sollen vor allem auch Sekundärdaten unterschiedlicher Quellen miteinander in Bezug setzen, um mögliche Wirkungszusammenhänge zu analysieren. Das bedarf eines weitaus komplizierteren methodischen Know-hows.

„Wir wissen heute viel mehr über Daten und Fakten als vor 15 Jahren, wo man Evaluierungen gebraucht hat, um überhaupt einen Eindruck zu bekommen. Heute braucht man Evaluation aus Pflicht oder um anspruchsvollere Fragen zu stellen.“

„In den letzten zehn Jahren wurde alles, was mit Bibliometrie zu tun hat, wirklich robuster, bedienungsfreundlicher, man kann es als einen Teil der meisten Evaluierungen einsetzen. Ansonsten ist meine Liebe zu qualitativen Methoden ungebrochen, bei anderen quantitativen Methoden besteht immer noch ein ziemlicher Gap zwischen Wunsch und Wirklichkeit.“

„Es ändert sich in der Technik einiges: Wie gehe ich mit Daten um, wie transparent werden diese bereitgestellt? Irgendwann wird der nationale Bereich zu den EU-Daten hinzukommen, dann wird eine relativ große Datenmenge frei zur Verfügung stehen, man wird Evaluationen nicht brauchen, weil das kann sich jeder suchen. Das heißt was für die Evaluation? Man kann Fragen adressieren, die über Programmgrenzen hinausweisen.“

„Die verschiedenen Stufen der Programmatik (frei zitiert nach Kantner) sind Themen, Kooperation, Kompetenz, Verwertungsrelevanz: Es geht nichts weg, es schichtet sich neu auf.“

- Das Europäische Förderprogramm „Horizon 2020“ organisiert die Forschungsförderung entlang großer Herausforderungen. Das bedeutet auch, dass die FTI-Politik-Evaluierung Bezug zu thematischen Feldern in der Gesellschaft und Sektorpolitik nimmt. Diese Felder verwenden jedoch unterschiedliche Sprachen und sind in unterschiedlichen institutionellen Settings verankert. Evaluatoren stehen hier vor der Herausforderung, sich neuen Analysefeldern und Kulturen zu öffnen.

„In dem Moment, wo wir aus unserer unmittelbaren forschungspolitischen Umgebung heraustreten, kommen wir drauf, wir sitzen alle im selben Wirkungsevaluierungsboot, können aber nicht miteinander reden, weil wir unterschiedliche Sprachen sprechen.“

- Wer diese Herausforderung anerkennt, stößt notwendigerweise auf neuen Lernbedarf. Manche der befragten Stakeholder optieren für begleitende Evaluation, die parallel zur Entwicklung und zum Start neuer Programme durchgeführt wird. So können laufend und über den gesamten Prozess hinweg Feedback-Schleifen genutzt werden, dies käme, wie oben bereits erwähnt,

einem Coaching-Prozess näher.

E. FTEVAL ALS KOMMUNIKATIONSPLATTFORM

Vom ersten Tag an wandte sich die fteval über ihren Newsletter, der später vom fteval Journal abgelöst wurde, an ein breiteres Publikum. Die unterschiedlichen Formate für Kommunikation und Vernetzung, die die fteval heute anbietet, werden von den Befragten durchwegs wegen ihrem interessanten Inhalt und ihrer Professionalität hoch geschätzt. Das betrifft kleine Veranstaltungen wie „fteval inside : fteval insight“ ebenso wie große internationale Konferenzen. Die befragten Stakeholder werden zu unzähligen Veranstaltungen eingeladen und sind dementsprechend selektiv, dennoch hatten sie alle die fteval Veranstaltungen als interessant und informativ in guter Erinnerung behalten. Neben Wissensvermittlung unterstützen die Veranstaltungen auch Vernetzung und „Community Building“, stärken das Vertrauen und helfen, eine gemeinsame Evaluierungskultur zu entwickeln, die die Basis für die gemeinsamen Standards ist.

„Was die Konferenz betrifft: Das war sehr aufwändig und ist organisatorisch an die Obergrenze gegangen. Wenn das Thema brennt, ist es gut. Wenn es eine Konferenz ist, weil es eine Konferenz ist, dann zählt es sich nicht aus; wenn kontrovers diskutiert werden kann, dann schon.“

„Die Konferenzen haben etwas Einzigartiges in Europa, weil sie schaffen, was andere Konferenzen nicht schaffen, sie sind nicht so stark wissenschaftlich orientiert und integrieren auch die Politik-Machenden, die Forschenden und die Beratungscommunity, teilweise auch aus dem Ausland. (...) Das Thema Evaluation würde aus reiner Forschungssicht auch nicht als Konferenzthema taugen.“

„Das fteval Journal hat eine lesbare Qualität und eine lesenswerte Qualität.“

„Das fteval Journal ist eines der wenigen Journale, wo man relativ sicher sein kann, dass das in Österreich von recht vielen Akteuren gelesen wird, die auch eine Bedeutung haben.“

Die Publikationen der fteval erhalten fast ebenso einhellige Zustimmung. Die meisten Befragten unterstreichen die Einmaligkeit und die Relevanz des Journals, das sie als anspruchsvoll und gleichzeitig praxisorientiert wahrnehmen. Alle Befragten – auch jene, die nicht in regelmäßigem Kontakt mit der fteval stehen – kennen es, manche lesen es regelmäßig, andere überfliegen mangels Zeit nur die Überschriften. Der Leser/innenkreis geht auch über die österreichische Grenze hinaus, manche der Autoren/innen wurden beispielsweise von Leser/innen aus Deutschland angeschrieben. Manche der Befragten sind jedoch kritischer und wünschen sich ein reaktives, flexibles Format, das eine an der Aktualität orientierte, zeitnahe Diskussion fördert, eher als elaborierte thematische Hefte. Das führt zur Frage des Formats an sich: Während die einen eindeutig bei der Papierversion bleiben möchten, würden andere diese aus Kostengründen und wegen des damit verbundenen Aufwands aufgeben. Wiederum andere plädieren für die Produktion sowohl eines gedruckten Journals als auch einer elektronischen Version, letztere soll nicht nur als pdf-file lesbar sein (dies ist heute schon der Fall), sondern sowohl für Smartphones als auch für e-books optimiert werden. Fraglos wird das Nebeneinander virtueller und realer Verbindungen bald den Auftritt, wahrscheinlich auch die Identität der fteval beeinflussen. Inter-

essanterweise hat der elektronische Newsletter, der seit 2013 an einen breiten Verteiler gesendet wird, vergleichsweise geringen Erfolg. Die Befragten erklären das durch den Schwall an elektronischer Information, der laufend in ihrer Mailbox landet.

Eine weitere Baustelle der Plattform ist ihr Internetauftritt. Die fteval verwaltet eine durchaus selten verfügbare Sammlung der meisten wenn nicht aller veröffentlichten Evaluierungsstudien eines Landes in einem bestimmten Politikfeld, nämlich der FTI-Politik in Österreich. Die meisten Befragten nutzen dieses Angebot oft oder dann und wann, und schätzen es, ähnlich einer Bibliothek auch zufällig auf Studien zu stoßen, die sie nicht aktiv gesucht hatten. Diese Datenbank, bislang schlicht als Download-Bereich organisiert, wird demnächst in ein Repositorium umgewandelt, mit detaillierten Suchfunktionen, klaren Identifikatoren und Textsuche. Wie schon weiter oben erwähnt, sind die Grenzen zwischen allgemeinen Studien und Evaluierungen nicht scharf gezogen. In diesem Zusammenhang plädieren viele der Befragten dafür, dass erstens einschlägige Auftragsstudien ebenso systematisch publiziert werden sollten wie es für Evaluierungen bereits üblich ist und dass diese zweitens im Repositorium erfasst sein sollen. Das klingt einfach, berührt jedoch die sensible Frage der gegenwärtigen Grenzen der Plattform beziehungsweise ihre Öffnung zu neuen Themen und in der Folge auch zu neuen Mitgliedern und Zielgruppen.

Dieses Spannungsfeld zwischen den Vorteilen einer klar fokussierten und hochqualitativen Plattform und der Notwendigkeit sowie dem Wert, sich auch mit anderen Bereichen in Verbindung zu setzen, seien dies regionale Akteure, Universitäten oder seien es thematische Bereiche wie beispielsweise Umwelt, Gesundheit oder Sicherheit, wurde in den Interviews immer wieder angeschnitten, ohne – und das liegt wohl in der Natur dieser Herausforderung – zu einem klaren Schluss zu kommen. Durchgängig wird jedoch begrüßt, dass sich die fteval nicht als "closed shop" versteht.

„Je enger die Community, desto enger auch die Plattform und desto eher entstehen auch informelle Auftraggeber-Auftragnehmer-Kontakte. Je offener das Ganze ist, desto eher verschwimmt es, auch die Wahrnehmung. (...) Die Öffnung nimmt auch schnell Kompetenz, es ist eh schwer, in der Evaluierung ein hohes Niveau zu halten. (...) Ich finde die Öffnungsstrategie der fteval wohltuend. Ich sehe auch umgekehrt die Limitierung, wer als Akteur infrage kommt.“

„Es gibt zwei große Abwesende: den Rechnungshof und das Finanzministerium. Die Plattform wirkt so, als ob sie die Komfortzone eher pflegt als infrage stellt. (...) Am Anfang war es sehr wichtig, um Standards aufzubauen.“

5. SCHLUSSFOLGERUNGEN UND HANDLUNGSFELDER

Aus diesem Spektrum an Erfahrungen und Bewertungen wollen wir einige Schlussfolgerungen und Handlungsfelder hervorheben, die uns, auch als Funktionsträger/innen der fteval, wichtig sind.

1. Die Plattform wird von allen Mitgliedern als hilfreich bewertet. Sie ermöglicht sowohl einen kontinuierlichen und nachhaltigen Austausch unter Mitgliedern und in der breiteren, auch internationalen Community als auch punktuelle Akzente wie Konferenzen und Trainings.
2. Die Evaluierungskultur der österreichischen FTI-Politik hat sich in den vergangenen 20 Jahren gut entwickelt. Die fteval hat hier eine wichtige Rolle gespielt. Heute stehen wir vor neuen Rahmenbedingungen – hier seien (i) knappe Budgets und der damit einhergehende Bedarf an Wirkungsmessung ebenso genannt wie (ii) die Verschränkung unterschiedlicher Politikfelder und die damit verbundenen Anforderungen an *policy intelligence*, aber auch (iii) Herausforderungen an quantitative Analysen neuerlich verfügbarer Datenmengen. Die Umfrage zeigt, dass die fteval und ihre Mitglieder Stärkefelder haben, besonders im Bereich der Programmevaluierung, dass es aber andere Bereiche gibt, die bislang unterbelichtet waren, insbesondere Evaluierung von Institutionen sowie eine transparente Vergabe und Veröffentlichung von ex-ante Untersuchungen zu Vorhaben der Forschungs- und Innovationspolitik. Diese Beobachtungen weisen darauf hin, dass die österreichische Evaluierungskultur und -praxis tendenziell bestätigend ist und nicht kritisch hinterfragt.
3. Ein spezielles Handlungsfeld sind *Terms of Reference (ToR)*, also die Pflichtenhefte der Evaluationen. Kurz gefasst gelten sie als systematisch überladen, gemessen an der grundsätzlichen Machbarkeit (was ist messbar?!) und am verfügbaren Budget. Darunter leiden nicht nur Evaluatoren, die den Eindruck haben, unter Wettbewerbsdruck „zu viel zu billig“ leisten zu müssen, es leiden auch der Fokus von Evaluierungen und damit der Auftraggeber. Während Evaluierungen in der Regel veröffentlicht werden, ist der Vergabeprozess inklusive ToR vertraulich. Aus den Interviews geht dennoch ein deutlicher Auftrag an die Plattform hervor, in diesem Gebiet den Austausch zwischen fteval Mitgliedern zu fördern und einschlägige Trainings zu organisieren, so dass zielgerichteter evaluiert werden kann.
4. Die Plattform lässt sich als Netzwerk verstehen, so ist auch die Frage nach Identität und Öffnung immer wieder neu zu stellen. Aus Abbildung 2 ist ersichtlich, dass die Mitgliederbasis sich kontinuierlich verbreiterte und dass die Gewichtung zwischen den vier Typen: Ministerien / Agenturen / Forschungsorganisationen / Beratungsunternehmen mehr oder weniger konstant blieb. Eine neue Ausrichtung gab es hingegen durch die verstärkte Einbindung von Akteuren aus dem Bereich wissenschaftlicher Forschung. Der nächste Schritt könnte – und sollte vermutlich – sein, in FTI-Politikevaluierung auch ein besseres Verständnis von Akteurskonstellationen einzubeziehen. Das betrifft die operative Ebene (SFB, DK, LBG, COMET, CDL⁸, Universitäten, Forschungseinrichtungen etc.) ebenso wie die Agentur-Ebene (FFG, AWS, ASI, AGES, UBA⁹,

8 Sonderforschungsbereiche, Doktoratskollegs, Ludwig Boltzmann-Gesellschaft, Christian Doppler Labors.

9 Österreichische Forschungsförderungsgesellschaft, Austria Wirtschaftsservice, Austrian Standards Institute, Österreichische Agentur für Ernährungssicherheit, Umweltbundesamt.

Patentamt etc.) und die Politik. Nicht alle der genannten Akteure sind Mitglieder der fteval und die Liste ist nicht vollständig. Sie soll nur illustrieren, dass es eine Vision der Plattform sein kann, den Weg zu bahnen, um – sei es experimentell – immer wieder aus einer neuen Perspektive die Wirkungsweise von FTI-Politik zu betrachten und diese nicht auf Förderungen (meist in Form von Zuschüssen und im Rahmen von Programmen) einzuschränken. Das vorliegende Heft, das unterschiedliche Intermediäre beleuchtet, ist bereits ein Schritt in diese Richtung. Ein weiterer Schritt, der seit den Anfängen der Plattform immer wieder gesetzt wird, ist die Öffnung für neue Mitglieder, bei hoher Achtsamkeit dahingehend, dass gerade wegen der unterschiedlichen Beziehungsmuster zwischen den Mitgliedern aufgrund ihrer Funktionen im Evaluierungssystem höchstmögliche Transparenz innerhalb der Plattform gewahrt wird.

5. Das letzte Handlungsfeld ist eine technische Modernisierung der Plattform in ihrer Publikationstätigkeit: Internetauftritt, elektronisch lesbare Publikationen, professionelles Bedienen sozialer Medien, Formate für Postings aktueller Diskussionsbeiträge zu FTI-Politik-Evaluierung, nicht zuletzt die Einrichtung eines Repositoriums für Evaluierungsstudien entsprechend der heute zugänglichen Qualitätsstandards. Das aktuelle Arbeitsprogramm der fteval hat die meisten dieser Punkte hier bereits aufgegriffen und die Umsetzung für das Jahr 2016 anberaunt.

Abschließend wollen wir an dieser Stelle allen Mitgliedsorganisationen für ihre Bereitschaft, an diesem Projekt mitzuwirken, danken. Ihr Engagement scheint uns ein lebendiger Beweis für das Commitment, das wir mit auf den Weg nehmen, um die genannten Herausforderungen gemeinsam aufzugreifen, zu sein.

AUTOR:INNEN

KATHARINA WARTA

Technopolis Group, Austria

E: warta@technopolis-group.com

STEFAN PHILIPP

fteval

E: philipp@fteval.at

EVALUATING THE PERFORMANCE OF INNOVATION INTERMEDIARIES: INSIGHTS FROM THE EXPERIENCE OF TUSCANY'S INNOVATION POLES

MARGHERITA RUSSO, ANNALISA CALOFFI, FEDERICA ROSSI, VALENTINA FIORELMONDO AND STEFANO GHINOI

ABSTRACT

With the growing importance of innovation intermediaries, particularly in the policy context, a need has emerged for appropriate instruments to evaluate their performance. The identification of appropriate performance indicators, however, has proved to be problematic. First, indicators are likely to influence the behavior of innovation intermediaries, not always in a desirable manner. Second, commonly used indicators focus on the immediate results achieved by the intermediaries, often disregarding the permanent behavioral changes that they can stimulate in their innovation system. Instead, we argue that the latter are particularly important for the evaluation of innovation intermediaries, whose success should be measured in terms of their ability to enable other organizations to improve their innovation capabilities. By focusing on an innovation policy intervention implemented by the Italian region of Tuscany in the period 2007-2013, we discuss the advantages and limitations of the indicators that have been set up by the regional government in order to evaluate the performance of innovation poles, a particular type of innovation intermediary, and discuss some feasible avenues for their improvement.

KEYWORDS:

innovation policy; innovation intermediaries; innovation poles; evaluation; technology transfer policies

1. INTRODUCTION

Intermediary organizations that support firm-level and collaborative innovation, often called 'innovation intermediaries', have gained increasing prominence in knowledge-intensive economies (Howells, 2006; Lazaric et al, 2008). They provide a range of knowledge-intensive services that include, among others, technology foresight and technology scouting, research and development (R&D) partnership formation, technical assistance in R&D projects, dissemination and commercialization of results, and technology transfer.

Since innovation intermediaries can facilitate knowledge exchange among organizations with different languages, cultures, decision-making

horizons, systems of incentives and objectives (Howells, 2006; Russo and Rossi, 2009; Caloffi et al, 2015), they can play a role in policies aimed at promoting innovation within local, regional and national systems (see e.g. Kauffeld-Monz and Fritsch, 2013). Examples of policy interventions that have funded organizations performing at least some innovation intermediary functions are the regional competitiveness poles in France, the Innovation Networks in Denmark, the Strategic Centres for Science, Technology and Innovation in Finland, the Catapult Centres in the UK (Russo et al. 2014). With the growing importance of innovation intermediaries, a need has emerged for appropriate instruments to evaluate their performance.

Often, policymakers define precise indicators for the ex-ante, ongoing and ex-post assessment of the activities carried out by innovation intermediaries. These indicators constitute an incentive system that alters the intermediaries' behavior: for example, they may induce the intermediaries to focus only on the requirements set by the policymakers, disregarding other objectives; or, given that indicators often focus on immediate outputs, intermediaries may be tempted to focus their efforts only on activities that produce immediate results and neglect activities that would yield results only over a longer time horizon.

By focusing on an innovation policy intervention implemented by the Italian region of Tuscany in 2007-2013, we discuss the indicators set up by the regional government in order to evaluate the performance of innovation poles (a particular type of innovation intermediary), and discuss some feasible avenues for improvement. Our focus is not simply on the immediate outputs of the intermediaries' activities, which would not have been achieved in their absence, but more importantly on the permanent behavioural changes (Georghiou, 1998; Hyvarinen and Rautiainen, 2007; Gok and Edler, 2012) that they have stimulated in the organizations they worked with: how they have helped organizations to change their way of innovating, rather than just how many services they have provided and how many firms they have supported.

The paper is organized as follows. Section 2 briefly introduces the main characteristics of the innovation poles implemented in Tuscany, and the performance indicators used by the policymaker. Section 3 discusses the limitations and implications of such indicators as perceived

by the policies' beneficiaries, whose views we have collected through a combination of semi-structured interviews, an online survey and a focus group. Section 4 describes the instruments and indicators that most poles used for their self-evaluation. The concluding section 5, combining the evidence base with theoretical insights, proposes a more comprehensive approach to evaluation in order to better capture the actual impact of innovation intermediaries on their regional innovation system.

2. THE POLICY PROGRAMME AND THE ASSESSMENT OF THE PERFORMANCE OF INNOVATION POLES

During 2011-2014, Tuscany's regional government funded twelve 'innovation poles', innovation intermediaries whose aim was to provide a range of knowledge-intensive services - including technology scouting, support in the preparation of project proposals, matchmaking - in order to encourage technology transfer and stimulate the innovation capabilities of regional small and medium-sized enterprises (SMEs). The final goal of this policy was to strengthen the regional innovation system, which includes a large number of SMEs that have relatively few connections with universities and other regional research hubs.

This intervention unfolded in two phases. In a first phase, Tuscany's regional government identified a set of twelve key technologies/applica-

tions, and it promoted the creation of an innovation pole for each of these. Having set the objectives and operational rules for the innovation poles, in 2010 the regional government launched a call for tender inviting organizations to submit proposals (individually or jointly) for the creation of twelve poles for a three-year period.

The 12 proposals selected for funding were each managed by a consortium, that included service providers, universities and firms. Overall, 49 different organizations were involved in the consortia managing the 12 innovation poles (some were involved in more than one pole), which, due to mergers and closures, dropped to 46 at the end of the three year period of the regional funding. Table 1 lists, for each innovation pole, its field (key technologies/applications), the number of managing organizations in the consortium and – since firms that intended to use an innovation pole's services would have to gain membership of that pole – the number of members at the start (30.6.2011) and end (30.6.2014) of the three year period. The poles received regional funds to carry out the following activities:

- marketing, to recruit new members to the pole, including scouting activities to encourage companies to demand knowledge-intensive services and to invest in innovation;
- participation in R&D projects at regional, national and European levels;
- management of the pole's open access infrastructures such as its laboratories;
- organization of knowledge transfer programmes, workshops and seminars to facilitate knowledge sharing and networking between members.

Technpole (acronym)	Key technologies/applications	N. of consortium participants	N. members as of 30.06.2011	N. members as of 30.06.2014
OPTOSCAN	Optoelectronics for manufacturing and aerospace	2	67	92
INNOPAPER	Paper	1	89	139
OTIR 2020	Fashion (textiles, apparel, leather, shoes, jewellery)	7	223	501
VITA	Life science	8	41	158
PIETRE	Marble	4	52	122
PENTA	Shipbuilding and maritime technology	5	225	352
POLIS	Technologies for sustainable cities	8	228	643
NANOXM	Nanotechnologies	6	70	128
CENTO	Furniture and interior design	6	177	322
PIERRE	Renewable energies and energy saving technology	13	120	368
POL012	Mechanics, particularly for automotive and transport	6	198	390
POLITER	ICT and robotics	13	195	697

Table 1: Consortium leaders, consortium participants, pole members

Source: our elaborations using data provided by Tuscany's Regional government

Regional funding to the innovation poles was conditional upon their achievement of a set of minimum performance targets. The tender stated that innovation poles would be assigned to one of three possible 'bands' depending on how many members they had at the time of their launch. Different performance targets were set for the different bands. Targets were defined as minimum thresholds with respect to: number of scout-companies, percentage increase in the number of member companies, number of services contracted and actually provided to firms, revenue. Table 2 shows, for each band, the minimum number of members required

at the start of the period, and the performance targets to be achieved over the three years. The rightmost column indicates the maximum funding that poles could claim from the regional government, which would only be released upon achievement of all their performance targets. The funding was given in two installments, one (up to 70%, provided that a bank guarantee had been presented) over the course of the three years, and the remaining amount at the end of the period (once the achievement of the performance targets had been ascertained).

	Criterion for allocation into bands:	Targets to be achieved within three years					Maximum funding that could be claimed from the Regional government
		scouting of new companies	% increase in members	n. knowledge-intensive services		minimum revenue	
	contracted			provided			
Band 1	> 160	160	50	80	40	500,000 €	800,000 €
Band 2	> 80	80	50	40	20	300,000 €	600,000 €
Band 3	> 40	40	50	20	10	150,000 €	400,000 €

Table 2: The innovation poles' classification into bands: criteria, performance targets and maximum funding that could be claimed
Source: our elaborations using data provided by Tuscany's Regional government

Table 3 shows the band each pole belonged to, and the funding they were granted. Six poles were in band 1, with more than 160 members at the time of their launch, while two were in band 2 and four were in band 3. Within the maximum limit provided for each band, the funding that could be granted to each pole was based on the budget they presented to the Region upon approval of their application.

Innovation pole	Band	Budgeted expenditure	Maximum funding that could be claimed from the Regional government	Granted funding
OPTOSCANA	3	505,056 €	400,000 €	280,000 €
INNOPAPER	2	757,543 €	600,000 €	420,000 €
OTIR 2020	1	1,007,000 €	800,000 €	560,000 €
VITA	3	595,000 €	400,000 €	280,000 €
PIETRE	3	-	-	-
PENTA	1	1,034,993 €	787,111 €	550,978 €
POLIS	1	1,007,000 €	800,000 €	560,000 €
NANOXM	3	503,500 €	400,000 €	280,000 €
CENTO	1	1,005,343 €	798,859 €	559,201 €
PIERRE	2	757,450 €	598,950 €	419,265 €
POLO12	1	1,007,000 €	764,391 €	535,074 €
POLITER	1	982,615 €	767,910 €	537,537 €

Note to table: Data are not available for the 'Pietre' innovation pole.

Table 3: Budgeted expenditure, maximum funding that could be claimed and funding actual-ly granted, by innovation pole Source: our elaborations using data provided by Tuscany's Regional government

For each of the performance indicators listed in Table 2, Table 4 summarizes the targets and the results achieved by each innovation pole, ordered by band. It is evident that all poles had reached their perfor-

mance targets by the end of the three years. Data (not reported here) show that some of them had already reached their targets three months after their launch.

Innovation pole	N. members	Scouting of new companies	Increase in member firms (above the minimum threshold for each band)	Services contracted	Services provided	Revenue
	<i>n.</i>	<i>n.</i>	%	<i>n.</i>	<i>n.</i>	€
Band 1	Minimum targets					
	160	160	100%	80	40	500,000
Results achieved by 30.06.2014						
OTIR 2020	501	278	201	108	93	2,085,734
PENTA	352	236	122	133	100	752,514
POLIS	643	274	306	111	88	690,087
CENTO	322	190	106	132	115	979,696
POLO12	390	249	146	251	267	1,838,604
POLITER	697	286	336	194	191	2,408,017
Band 2	Minimum targets					
	80	80	100%	40	20	300,000
Results achieved by 30.06.2014						
INNOPAPER	139	94	74	446	455	731,209
PIERRE	368	120	360	85	64	640,655
Band 3	Minimum targets					
	40	40	100%	20	10	150,000
Results achieved by 30.06.2014						
OPTOSCANA	92	56	130	27	42	684,431
VITA	158	73	295	34	31	246,382
PIETRE*	122	81	202	18	18	1,487,107
NANOXM	128	44	155	43	25	531,845

*Forecasted as of 30.06.2014

Table 4: Performance targets and their achievement

Source: our elaborations using data provided by the innovation poles

Even a quick glance at the poles' performance poles suggests that the indicators and targets set by the regional government might have created some incentives that were not entirely desirable. First, the fact that most poles had reached their targets very rapidly suggests that such targets were probably too low. Innovation poles may have quickly implemented actions aimed at achieving them, possibly regardless of whether such actions were appropriate at such an early stage. Second, since the poles' performance was evaluated only on the basis of the achievement of these targets, most of them stopped reporting (or even stopped systematically collecting) any information on the activities they had carried out after their targets had been achieved. Therefore, our knowledge about the poles' activities could be incomplete.

The lack of satisfaction with current indicators was shared by regional officers (particularly during the debate on the design of the new policy on technological districts) and by the innovation poles themselves. When asked whether the poles' performance could have been more effectively measured by different indicators, nine out of twelve of the managing consortia's leaders agreed. The indicators adopted by the regional government, in fact, collected quantitative information that was not sufficient to capture the actual activities performed by the poles and the role of their managing consortia in supporting the region's innovation system.

3. LIMITATIONS AND IMPLICATIONS OF CURRENT INDICATORS

Our analysis builds upon several data sources, which the authors assembled as part of a research team engaged in the analysis of the policy programme. First, the team collected secondary information either in the public domain (the poles' websites) or provided by Tuscany's regional government (the six-monthly activity reports submitted by each pole to the regional government). Second, 27 semi-structured interviews were carried out between March and May 2014, with representatives of the organizations managing the innovation poles, of some member companies and of local business associations, and with an employee of the regional government office managing this programme. These interviews, however, were not aimed primarily at exploring the limits and implications of the poles' performance evaluation process. Therefore, we supplemented them with two additional information sources, an online survey and a focus group¹. The online survey aimed to capture the views of the poles' managing organizations on the monitoring and evaluation activities they had to comply with over the course of the three years. In particular, the managing organizations were asked to evaluate the effectiveness and efficiency of the indicators used by the regional government to assess the poles' performance, to explain whether each innovation pole had engaged in processes of self-assessment, and if so on the basis of what indicators and processes. Only 14 (30%) out of the 46 managing organization responded to the survey, 12 of these being the consortium leaders for the 12 poles².

The focus group involved representatives from nine innovation poles, besides the research group.

The results of the survey and the focus group discussion highlighted four main critiques to indicators used for policy evaluation:

- some indicators were deemed unsuitable to evaluate the poles' performance in several key activities;
- some important activities were not measured and therefore not used for performance evaluation;
- the poles' sectoral and technological specificities were not accounted for;
- the performance evaluation system did not encourage the poles to collect data accurately and systematically, and it imposed too many administrative requirements on the poles and the users of the services they provided.

3.1 INADEQUATE INDICATORS

As shown in Table 2, the performance targets, and the relative indicators to measure them, focused on two main activities: the scouting of new companies, aimed at recruiting new members, and the provision of knowledge-intensive services to poles' members. Through the online survey, the leaders of the managing consortia identified a number of limitations of the indicators used.

Concerning the provision of knowledge-intensive services, several poles claimed that the regional government had adopted a too narrow definition of what services could be included in the computation of the number of services contracted and actually provided, and of their revenue. In particular, any services that were purchased by the pole's members where the innovation pole operated as an intermediary rather than a direct service provider, were excluded. Several poles suggested that this intermediating activity should have been measured, since it was in line with the poles' general objective to increase the demand for knowledge-intensive services on the part of regional firms. In particular, some services had been provided by organizations that were members of the poles, but which were not part of their managing consortia: in this case, the poles effectively enabled connections between their members, promoting both demand and supply of knowledge-intensive services. Moreover, the services provided were not differentiated according to their quality and impact; for example, research provided as part of a large scale research project would have much greater 'knowledge content' than an ordinary prototyping service. Differentiating between services with different knowledge intensity would have allowed the policymaker to apply different weights to different services, and to better capture their different impacts.

Numerous remarks also concerned the approach used to evaluate the performance of scouting activities. The indicators only focused on the number of companies that were contacted and those that subsequently became members. However, evidence from the focus group suggests that 'scouting' included a complex set of activities³, which required more sophisticated indicators. The consortium leaders suggested that these indicators not only should have been differentiated according to the different nature of the companies targeted in different sectors (a point to which we return later in this section), but they should have taken into account the scouting activities' medium-term impacts, focusing on the presence of follow-up activities after each initial contact (indicating that the contact had been productive). Examples of follow-up activities that could have been measured are: number of contracts that the scouted companies subsequently signed with the pole (or thanks to its intermediation); number of project proposals that the scouted companies subsequently presented together with the pole; whether the scouted

1 The first email inviting the managing organizations to take part in the survey was sent on March 27th, 2015, followed by two recalls sent to non-respondents only. The focus group was held in May 2015 in Florence.

2 We focused our analysis on the 12 completed questionnaires received from the 12 consortium leaders.

3 The discussion highlighted that most innovation poles had structured their scouting activities in two main phases. In the first phase, the poles' objective was to attract members, in order to reach their performance targets. This phase often relied on approaching the managing organizations' pre-existing networks, mapping their potential members and recruiting them to the innovation pole. Once the target had been achieved, the poles launched a second phase in which the scouting activities were aimed at better

companies had subsequently participated in workshops of other events organized by the pole; how many follow up meetings were scheduled with these companies.

3.2 MISSING ACTIVITIES

To understand whether the poles had been effective in supporting innovation and technology transfer, further activities should have been taken into account. Our survey showed that eleven out of twelve poles performed activities beyond those measured by the indicators, in particular they engaged in (i) preparation of project proposals, (ii) networking and (iii) other initiatives.

Most poles provided support to companies in the preparation of project proposals to be submitted to public tenders (European, national, regional, etc.); however, they usually kept a record of this activity only if the proposal had subsequently been funded. Indicators that the consortium leaders suggested in order to measure this activity include the number of project proposals submitted and the share of submitted applications that were successful.

The poles were very active in the organization of events aimed at promoting networking between their members, potential suppliers and research laboratories. Among these, the respondents mentioned matchmaking events, seminars, workshops, study visits, training events, attendance to fairs and exhibitions. The proposed indicators included the number of events organized, by type, and the number of companies that attended each type of event (possibly also as share of all member companies, and as share of companies present in that sector), as well as indicators capturing the follow-up activities to each event (further contacts, projects, etc.)

Consortium leaders also mentioned varied activities that were, to some extent, linked to the poles' intermediation activities: attraction of investment in the local area, replication of the technology transfer model in other regions, networking between public agencies and SMEs, participation in international collaborations and membership in national and European clusters as well as agreements with public agencies and trade and professional bodies. As we already noted in section 3.1, activities in which the poles played the role of intermediaries (rather than the role of direct service providers) were not considered in performance evaluation, but the poles viewed them as important components of their overall engagement.

3.3. LACK OF SECTORAL/TECHNOLOGICAL DIFFERENTIATION

A cross-cutting theme that emerged from the survey and the focus group, as well as from the initial interviews, is the need to account for the fact that each pole deals with very different sectors, where organizations differ in terms of size, research intensity, technologies, modes of innovation, nature, localization of the demand and of the value chain, relationships with clients, suppliers and competitors. Two consortium leaders explicitly stated the need to distinguish between poles that operate in traditional sectors and poles that operate in high tech sectors based on key enabling technologies. The issue of the specificities of individual poles has emerged several times both in the context of the general dis-

cussion of the most appropriate indicators for the poles' evaluation, and in the context of more specific themes like the evaluation of scouting activities. Here, it has been suggested that performance indicators should take into account the different characteristics of the scouted companies (such as size or specialization), the different characteristics of the sector to which they belong (such as the number of companies present in it), and the different characteristics of member organizations (such as turnover, employees, production units, etc.). No suggestion was advanced on what might be good metrics to assess those differences.

3.4. INADEQUATE DATA COLLECTION TOOLS

To support the evaluation process, it would have been important to encourage the poles to systematically collect information on the activities they performed. For example, it would have been useful if each pole had collected in a relational database the details of the companies they contacted, or that they intended to contact, and of the outcomes of the scouting activities with these companies. Not only this would have supported monitoring and evaluation, but it would also have created a precious knowledge base for the regional government and the overall poles system. While some poles, in order to support their scouting activities, had systematically mapped the companies in their target sector, this exercise was limited to a few poles, the information collected was very diverse, and so were the tools used (in some cases, data were collected only in paper form). The focus group participants suggested that greater coordination and sharing in the process of information collection within and between poles would have been useful. This could have been supported by a preliminary effort on the part of the policymaker to provide clear guidelines for the collection of information, clear information about which data would be used in the construction of indicators, and possibly some information collection tools such as pre-structured databases ready to be filled in. But, *ex ante*, the policymaker was not fully prepared to cope with the cascades of novelties emerging from the activities of the new innovation intermediaries they were funding.

At the same time, the focus group participants suggested that data collection instruments should be easy to use and avoid imposing a too heavy administrative burden on the poles and, more importantly, on the companies they dealt with. Some of the administrative requirements imposed by the regional government were, indeed, quite cumbersome⁴.

4. THE INNOVATION POLES' SELF-EVALUATION PROCESSES

Evidence from the online survey suggests that nine out of twelve innovation poles engaged in self-evaluation exercises based on richer sources of information than the set of indicators used by the policymaker, in order to better understand their own performance and to support their planning and strategy development. Of these nine, one pole used only qualitative information, one used only quantitative indicators, and the other seven used a combination of qualitative and quantitative information.

Concerning quantitative indicators, the survey respondents confirmed that they counted some important activities which were not included in the performance indicators used by the regional government: they counted events like seminars, focus groups, workshops aimed at sharing information with companies, and matchmaking events between companies and research centres; they also counted any further collaborations that were initiated as a consequence of the networking activity enabled by the pole. One innovation pole suggested that the number of interactions with organizations outside the region could indicate the pole's ability to expand the geographical scope of its activities. Some poles monitored the preparation of project proposals, including those that were unsuccessful, collecting information about the number of project applications that were submitted, the number of companies and research organizations involved in these project applications, as well as the amount of public funds that the organizations involved in these project applications (particularly companies) were able to accrue. Another variable that several survey respondents indicated as being very important is the participation in national and international clusters and in other international and national collaborations. While several poles used the number of such collaboration agreements as an indicator of their own performance, many warned that this information should be complemented by information about their actual effectiveness and about the follow-up activities realized.

Many poles also evaluated their own performance on the basis of qualitative information about: their members' degree of satisfaction, the members' ease of access to laboratories, the quality of the relationships that were created between companies and research organizations, the increase in visibility and in recognition of the poles' activities, the poles' impact on other sectors of the regional economy, the activities performed by the temporary managers and the outcomes of the matchmaking events.

Most poles suggested that these variables, both qualitative and quantitative, should have been taken into account in the performance evaluation process, in order to produce a more comprehensive overview of the poles' activities. Collecting qualitative information would also have allowed the poles and the policymaker to gain insight into the sectors' technological trends, would have facilitated technology foresight and the understanding of the poles' competitive positioning within and outside the region.

5. CONCLUSION: TOWARDS A BETTER APPROACH TO ASSESSING THE PERFORMANCE OF INNOVATION INTERMEDIARIES

In order to comprehensively evaluate the poles' performance in light of their objectives to encourage technology transfer and stimulate the innovation capabilities of SMEs, ultimately strengthening the competitiveness of the region's innovation system, it would be important to adopt a broader perspective, beyond the immediate outputs of the poles'

activities which were captured by the indicators adopted by the regional government. It is worth noting that, in Italy, other regions adopting innovation policies similar to Tuscany were using a similar set of indicators to assess performance of the innovation intermediaries. With regard to our empirical analysis on Tuscany, we observed that the poles' own criticism of the policymaker's approach to evaluation focused on the lack of attention for certain types of activities and for certain types of "follow-up" activities that generated medium-term outputs. But neither the policymaker nor the poles perceived the importance to capture the permanent behavioural changes (Georghiou, 1998; Hyvarinen and Rautiainen, 2007; Gok and Edler, 2012) that the poles stimulated in the organizations they engaged with: how they changed their way of innovating, rather than just how many more knowledge-intensive services they demanded. In order to capture these effects, the focus should have been on the learning processes that had been activated thanks to the programme, which might have led to changes in policies, structures and strategies, both for the poles themselves and for the organizations they engaged with.

On the one hand, it would have been useful to explore the skills and competences possessed by the poles themselves, and their evolution over time, to analyse how these changed, and whether any changes occurred in response to contingent events, or whether they resulted from a process of learning from experience. For this purpose, qualitative information could have been collected by asking the poles' managing organizations questions such as: What skills and competences did the poles (and their managing organizations) possess at the start of the programme? Did they change over time? If so, how? What processes stimulated these changes? Did the poles change their operations, their governance, their business strategies?

In our survey, the consortium leaders agreed that they had derived benefits from their involvement with one or more poles, not just in financial terms but also in terms of relationships, skills and visibility. Thanks to their involvement with the poles, they increased their number of relationships with companies, not only in their own sector of expertise but also in others. Their knowledge of these companies improved, allowing them to gain more detailed insight into the needs and weaknesses of these companies and of the sector overall. They also improved their technology transfer skills, allowing them to provide a broader range of services and giving them the ability to network with international partners. Finally, their visibility increased, allowing them to attract better quality human resources and enjoy higher reputation with companies and other potential partners.

On the other hand, it would have been important to also analyze the changes induced in the companies that engaged with the poles. An analysis of the evolution of their demand for knowledge-intensive services could have generated valuable information about the effectiveness of the policies and the ability of the poles to achieve the desired goals. An ongoing survey commissioned by Tuscany's regional government to the regional research agency IRPET is exploring this domain by collecting information on a sample of companies members of the innovation poles. The focus is on questions such as: why did the company demand services from the pole rather than from competing providers (if they existed)? Did the nature of the services demanded change over time? Did the value of the services demanded change over time? Did the frequency with which services were demanded change over time?

Other useful information could have been collected about the changes in the members' networks of relationships as a consequence of their involvement with the poles; as well as about changes in their innovation activities, in terms of the nature and value of the research project proposals they submitted and for which they received funding, of the other types of collaborative projects they engaged in, of the nature and types of investments in innovation they made.

The following table summarizes the possible indicators for the general evaluation of innovation intermediaries' performance which have emerged from our discussions with the poles' managing organizations.

We have categorized these indicators according to whether they capture the immediate outputs of the innovation intermediaries' activities (column A), the medium-term follow-on impacts of these actions (column B), or the long-term behavioural changes induced by the innovation intermediaries' activities (column C). The indicators are also categorized according to the type of activity they refer to, considering four types which reflect those introduced in section 2: (i) scouting of new companies, (ii) provision of knowledge-intensive services, (iii) preparation of project proposals, (iv) networking and (v) other initiatives.

Type of activity	A. Immediate outputs	B. Follow on impacts	C. Long term behavioural changes
Scouting of new companies	Number of companies scouted Share of large companies and of SMEs	Share of companies - that attended at least one event organized by the innovation intermediary (round tables, workshops, seminars etc.) - with which at least one follow-up meeting has taken place - with which at least one follow-up activity has been carried out Number and value of follow-up activities carried out with the companies involved in scouting, by type (types of follow-up activities can include: successful project applications, unsuccessful project applications, inter-organizational consortia, other projects)	Qualitative analysis - of the objectives of the innovation intermediary's scouting activities and whether they changed over time - of whether and how the scouting activities implemented by the innovation intermediary changed over time (types of activities, skills required, procedures and approaches used) - of what the innovation intermediary's managing organizations learned from their experience of the scouting - of what did scouted companies learn from their experience of the scouting
Provision of knowledge-intensive services	Number and value of knowledge-intensive services provided to member companies: - by any of the innovation intermediary's managing organizations - by other actors (intermediated by the innovation intermediary)	Number and value of knowledge-intensive services provided to: - companies by any of the innovation intermediary's managing organizations (by type of service and over time) - member companies by other actors (intermediated by the innovation intermediary), by type of service and over time	Qualitative analysis (questionnaire to member companies) to investigate why the companies have used that innovation intermediary Qualitative analysis - of what the innovation intermediary's managing organizations learned from their experience of providing knowledge-intensive services - of what did (a sample of) scouted companies learn from their experience of receiving knowledge-intensive services Analysis of the evolution in the services demanded, in terms of nature, value and frequency
Project proposals	Number and value of - project proposals (submitted and funded) presented to national tenders international tenders	Number and value of - project proposals (submitted and funded) presented by member companies with the support of the innovation intermediary	Learning outcomes Qualitative analysis of what - the innovation intermediary's managing organizations learned from their experience of supporting the preparation of project proposals - member companies learned from their experience of preparing project proposals with the support of the innovation intermediary
Networking	Member companies: - percentage increase - as a share of all companies in the targeted sector Number of events held, by type (matchmaking events, seminars, workshops, study tours, training events, etc.) Number of companies participating in events held, by type (matchmaking events, seminars, workshops, study tours, training events, etc.)	Share of member companies that - have demanded knowledge-intensive services (whether offered by the innovation intermediary's managing consortium or by others, but with the intermediation of the innovation pole) - have participated in events organized by the innovation intermediary - have activated collaborations with other member companies Number and value of follow-up activities jointly carried out by member companies (types of follow-up activities can include: successful project applications, unsuccessful project applications, inter-organizational consortia, other projects)	Learning outcomes Qualitative analysis of what - the innovation intermediary's managing organizations learned from their experience of organizing networking events - member companies learned from their experience of participating in events Changes in the networking activities of the member companies
Other initiatives	Description of the other initiatives carried out by the innovation intermediary (where possible, indicating number and value)	Description of the follow-up actions emerging from the other initiatives carried out by the innovation intermediary	Learning processes resulting from the other initiatives carried out by the innovation intermediary

Table 5: Proposed indicators to capture the poles' performance

Which of these indicators to use, and in what circumstances? In order to identify the appropriate approach to assess the performance of innovation intermediaries, it is important to distinguish between two conceptually different objectives that performance indicators could fulfill: (i) the evaluation of the public investment, and (ii) the evaluation of the performance of the innovation intermediary and of its sustainability.

Indicators used to support the allocation of public funds should privilege clarity, ease of construction and use: they should be few in number, simple, linked to the intermediaries' key activities, and they could be expressed in terms of minimum thresholds to be achieved. When evaluation processes are based exclusively upon indicators, particularly quantitative ones, there is always a risk that such indicators might be pursued for their own sake, rather than as instruments to achieve broader objectives, and that targeting one's behaviour to these indicators might prevent or hamper the achievement of these broader objectives.⁵ Therefore, indicators should have a direct link to one or more explicitly defined policy objectives, and their construction should always be accompanied by a reflection on what behavioural incentives they might generate, in order to avoid the creation of incentives that are misaligned with the broader objectives that the policy intervention intends to serve. Based on the our empirical evidence, we argue that, in order to evaluate how public funds were used, the regional government could have adopted a broader range of indicators, including those on "immediate outputs" and "follow on impacts" reported in columns A and B of Table 5.

Indicators aimed at returning a comprehensive evaluation of the performance of the innovation intermediaries (to evaluate the extent to which they encourage technology transfer and stimulate the innovation capabilities of SMEs, ultimately strengthening the competitiveness of the region's innovation system) should be comprehensive, detailed and expressed in terms of actual indicator values and processes enhanced, in order to provide a full picture of their success and sustainability. These indicators should also include the mainly qualitative information proposed in column C of table 5 on "long term behavioural changes". Furthermore, self-evaluation should be encouraged, as it not only allows any organization to identify its strengths and weaknesses, and then to implement actions to address them, but also because the very act of engaging in self-assessment implies reflection on their objectives and instruments in the short and long term, which may not happen at all in the absence of this action. In the context of policies supporting innovation intermediaries, the self-evaluation process could be formalized as part of the evaluation of the project proposals that are candidate for funding: that is, applicants who propose to set up publicly funded innovation intermediaries should explain how they are going to carry out the self-evaluation of their activities and introduce the qualitative and quantitative indicators that will be used in this process, the tools that will support the evaluation, how these indicators would be compiled and updated, and the frequency with which the self-evaluation would be car-

ried out. The suitability of self-evaluation to capture the relevant aspects of the structure's performance, would become part of the overall ex ante evaluation of the proposal in order to decide eligibility.

In the next steps of our research we will use some of the indicators discussed above to analyze the behavioral changes generated by the innovation poles.

REFERENCES

- Caloffi, A., Rossi, F., Russo, M.** (2015). The Emergence of Intermediary Organizations: A Network-based Approach to the Design of Innovation Policies. In Geyer R., Cairney, P. (Eds) *Handbook on Complexity and Public Policy*, Cheltenham: Edward Elgar.
- Georghiou, L.** (1998). Issues in the evaluation of innovation and technology policy. *evaluation*, 4(1), 37-51.
- Gök, A., & Edler, J.** (2012). The use of behavioural additionality evaluation in innovation policy making. *Research Evaluation*, 21(4), 306-318.
- Howells, J.** (2006). Intermediation and the role of intermediaries in innovation. *Research policy*, 35(5), 715-728.
- Hyvärinen, J., & Rautiainen, A. M.** (2007). Measuring additionality and systemic impacts of public research and development funding—the case of TEKES, Finland. *Research evaluation*, 16(3), 205-215.
- Kauffeld-Monz, M., & Fritsch, M.** (2013). Who are the knowledge brokers in regional systems of innovation? A multi-actor network analysis. *Regional Studies*, 47(5), 669-685.
- Lazarc, N., Longhi, C., & Thomas, C.** (2008). Gatekeepers of knowledge versus platforms of knowledge: from potential to realized absorptive capacity. *Regional Studies*, 42(6), 837-852.
- Russo, M., & Rossi, F.** (2009). Cooperation Networks and Innovation A Complex Systems Perspective to the Analysis and Evaluation of a Regional Innovation Policy Programme. *Evaluation*, 15(1), 75-99.
- Russo, M., et al.** (2014). Politiche a sostegno del sistema di ricerca e sviluppo in Danimarca, Finlandia, Francia, Germania, Italia, Spagna e Svezia, DEMB Working Paper Series, n. 45 [available at http://merlino.unimo.it/campusone/web_dep/wpdebm/0045.pdf]

5

As we have stated earlier, simply using the number of members as a measure of performance might have lead poles to undertake actions aimed at recruiting the largest possible amount of members without sufficiently investing in the quality of these relationships. Or, measuring the performance of the pole on the basis of the overall revenue accrued through services provision might have encouraged the poles to adopt strategies aimed at maximising this variable (for example, providing many services of small value, or providing a few very expensive services to a few clients) but these strategies may not necessarily serve the broader objective to support innovation and competitiveness of the firms operating in the targeted sector.

AUTHORS

MARGHERITA RUSSO

Department of Economics "Marco Biagi", University of Modena and Reggio Emilia, Italy

E: margherita.russo@unimore.it

ANNALISA CALOFFI

Department of Economics and management, University of Padua, Italy

E: annalisa.caloffi@unipd.it

FEDERICA ROSSI

Department of Management, Birkbeck College, University of London, United Kingdom

E: f.rossi@bbk.ac.uk

VALENTINA FIORELMONDO

Department of Economics "Marco Biagi", University of Modena and Reggio Emilia, Italy

E: valentina.fiordelmondo@unimore.it

STEFANO GHINOI

Department of Economics "Marco Biagi", University of Modena and Reggio Emilia, Italy

E: stefano.ghinoi@unimore.it

EVALUATING CLUSTER INITIATIVES: AGENCY, ORGANISATION, FUNCTIONALITY, PERFORMANCE

MICHAEL ROTHGANG AND BERNHARD LAGEMAN

CLUSTER INITIATIVES AS OBJECTS OF EVALUATION

Intermediary organisations are an integral part of any developed market economy. They fulfil essential tasks in the organisation of markets and the execution of government policies in market environments. While the attention of economic theory is mainly concentrated on the operation and interplay of markets and the relationships between market and state, intermediaries of all kinds play an essential mediating role without which neither markets nor the state could function smoothly.

Unlike, for instance, general fiscal promotion of firm R&D, cluster policy requires mediators that are able to effectively connect policy administrators with the cluster actors, that is, firms, research organisations (below used for universities and research institutes) and service providers, who are all members of the cluster innovation network. While indeed it is true that “clusters” are the ultimate target of cluster policy, it is clear that a cluster is an abstract, analytically created unit which cannot speak or act for itself and cannot engage in legal transactions. Thus, cluster policy needs agents who are capable of articulating the cluster population’s needs, goals and claims.

The required **mediation task** can either be assigned, for example, to prominent personalities from public life or civil servants – “one-person **cluster initiatives**” – or to more complex organisational structures, i.e. fully developed, formalised cluster initiatives. These **cluster initiatives** develop in a more or less democratic founding process which involves a part of the cluster population (firms, research institutions, public agencies). The first mode of mediation establishment is typical for top-down approaches of “cluster building”¹, the second mode is characteristic for bottom-up approaches. In practice, the development of cluster initiatives often is characterized by a mix of both development modes. In this contribution, we concentrate on formalised cluster initiatives disposing at a

mission statement or a strategy document, a cluster management and one or several cluster boards.

The terms “cluster” and “cluster organisation” denominate two fundamentally different objects of investigation: According to a definition by Porter, “A cluster is a geographically proximate group of interconnected companies and associated institutions in a particular field, linked by commonalities and complementarities. The geographic scope of a cluster can range from a single city or state to a country or even a network of neighboring countries.” (Porter 1998, p. 199). According to the original version of the “Cluster Initiative Greenbook” (Sölvell, Lindquist, Ketels 2003, p. 9) that we adopt in the following discussion, “Cluster initiatives are organized efforts to increase the growth and competitiveness of clusters within a region, involving cluster firms, government and/or the research community. Thus, “clusters” are agglomerations of firms and other organisations which feature certain attributes, whereas “cluster initiatives” are organized efforts – in most cases formal organisations – which aim at creating or furthering a cluster of firms and associated institutions.

With the rapid dissemination of cluster policies in the industrialised world since 1990, a whole range of such intermediaries operating between government authorities and cluster populations made its appearance. The Global Cluster Initiative Survey 2012 (European Commission 2012, pp. 9-11) which analysed the survey responses of 435 cluster initiatives, found that 36% of the answering initiatives were initiated as “a response to a public call /policy programme” and 39% as “a private sector initiative”.² More than 50% of the initiatives performed “time-limited activities for specific purposes involving only certain sub-groups of participants” (p. 11), the rest offered basic infrastructural services that were mainly open to everyone. The findings presented in Lindquist et al. (2013, p. 14)³ show that the vast majority (81%) of the cluster initiatives were launched after 2000, 59% were established in 2007 or later. In this context, Sölvell (2009) and Andersson et al. (2004) present systematic analyses of cluster initiatives.

1 In cluster research, the question if the public administration could be able to create an emerging cluster is, to say the least, highly controversial. Nevertheless, the colloquial phrase that one is about “to create a cluster” is often heard on the floors in the relevant public institutions, at least in Germany and probably not only there.

2 As the delimitation between both options is not quite clear in the source, there may be overlaps. 25% voted for the answer “other” which may be interpreted in different ways.

3 Same data collection as first source, but different details referring to the basic population, in this case 356 respondents.

In most research works on cluster policy, clusters (a particular type of economic agglomerations) as the “final” object of political intervention are in the focus of the analysis, whereas cluster initiatives (organisations marked by the intended purpose to create or further a cluster) as object of research per se are given only incidental coverage. This seems to be a shortcoming because these initiatives are one central element of cluster policy that are important for understanding its success and failure. Other relevant factors are, for instance, the general framework of a cluster program, extent and modes of funding, and the evolution of the relevant markets. On a closer look, there are construction problems inherent in the building and evolution of these intermediary organisations which are well capable to influence the success of cluster policy. Therefore, evaluations of cluster programmes should regard the initiatives as important success (or failure) factors of the programmes and hence critically scrutinize them.

Our contribution brings into focus the component of cluster programme evaluations that deals with cluster initiatives. We shall expand on the wider horizon of cluster programme evaluations only at the end of the discussion, when it comes to have a look on the “performance” of the cluster initiatives. The discussion is inspired by the common structural and process patterns of technology cluster initiatives which mostly can be found in public cluster programmes.⁴

The work is based on an analysis of the disposable literature, of generally accessible internet information and, last but not least, the authors’ experiences gathered in the course of the evaluation of the German “Leading-Edge Clusters Competition” (LECC) in the time period from 2009 to 2014 (Rothgang et al. 2014, 2015). Examples and experiences from the LECC are used to illustrate the central issues of an evaluation of cluster initiatives and the methods and indicators which could be applied in this context. Theoretical foundations are drawn from cluster research, institutional economics and organisational theory.

The four central dimensions of the analysis of cluster initiatives which are discussed below are agency, organisational construction, functionality and performance. For each of these dimensions, we discuss their role in the evaluation, the theoretical foundation, practical results and observations, and questions that should be addressed in the evaluation of the cluster initiatives. On that background, we analyse the relationships between these dimensions and discuss the larger context of cluster programmes in which the cluster initiatives operate.

EVALUATING CLUSTER INITIATIVES

AGENCY (A)

Evaluators who have the task to assess any cluster policy programme will be confronted in the first instance with cluster initiatives which claim to represent the respective clusters and to be entitled to speak for them. This claim of legitimacy will normally seem to be a matter which

does not need to be critically scrutinised in the eyes of most stakeholder involved. These stakeholders are the initiative’s founders, speakers of the participating cluster firms, local administrators providing public support as well as programme administrators. However, evaluators cannot avoid to analyse the initiative’s context and origins. The features of the respective cluster as reference unit and the embeddedness of the initiative in the cluster population have to be analysed in order to develop a deeper understanding of the functioning of cluster organisations. Core aspects of agency of cluster initiatives to be analysed are the degree of representation with regard to the cluster population as well as the ways how interests and actions of the initiative’s organisers and cluster firms are interwoven.

From a theoretical point of view, agency is a general challenge that organisations of all kinds have to meet with regard to their internal structures and, as soon as there are integrated inter-organisational networks, with regard to their external relationships. A cluster initiative, a special case of an inter-organisational network, is acting as an agent on behalf of its member organisations, mostly cluster firms and research organisations, education and training organisations but also other intermediaries such as chambers of commerce or public agencies. These member organisations (cluster actors) act as principals who are interested in the agent’s correct fulfilment of the common tasks of the initiative which were defined by their relevant decision-making body. At the same time, member organisations have a keen interest to receive a range of benefits from the cluster initiative’s activities which will normally differ depending on the type of organisation involved (e.g. enterprises and research organisations). The cluster management being the gatekeeper between cluster organisers and cluster actors, an information asymmetry between both seems to be unavoidable. In this case the problems of hidden information and hidden action dealt with in institutional economics could become relevant. Alike, the critical problem of an “equal” (of fair) distribution of cluster initiative’s benefits between cluster actors, e.g. in the formulation of cluster strategies and the specification of common research priorities in technology clusters, has to be taken into consideration.

A closer look on cluster initiatives which were operating under ambitious cluster programmes shows that the degree of the clusters’ population involvement in the respective initiatives differs markedly (figure 1). Many of the cluster firms and research organisations cooperate by inscribing themselves as formal members, some of them take an active, may be leading part in the initiative’s activities, others aim at a formal membership which is not linked to any deeper engagement. Finally, there are always some cluster firms that seem to be with regard to their product portfolios “born” members of the cluster initiative but who are standing aloof when it comes to the cluster initiative’s formation.

In the case of the LECC, the numbers of “cluster partners” according to the cluster managements’ lists of members varied considerably, from a minimum of 31 (Forum Organic Electronics) to a maximum of 349 (MicroTEC Südwest) (Rothgang et al. 2014, p. 51). The total number of listed cluster partners in the 15 “leading-edge clusters” of the three consecuti-

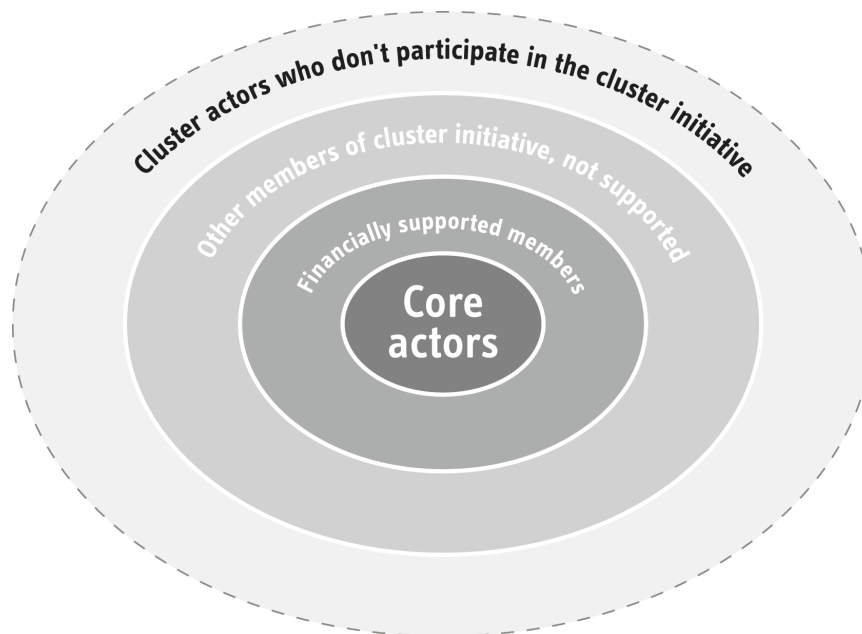


Figure 1: Varying Degrees of Cluster Actors' Engagement in the Cluster Initiative
Source: Own depiction, modified version of Rothgang et al. 2014, p. 50.

ve competition rounds amounted to 2.349 (on average 157) in December 2012, the number of formal members slightly varying over time because of withdrawals and entries.

Questions that should be answered in the evaluation of the cluster initiatives with regard to agency are:

- What is the general context of the cluster initiative's foundation (industry, technology field, competing cluster initiatives in the cluster area, influence of government agencies on the initiative's formation, strengths and weaknesses of the region's infrastructure?
- How did the founding process proceed (identity and role of initiator's and facilitators)?
- Which firms and research organisations form the cluster population according an empirical analysis of the assumed cluster region?
- Which of the firms and research organisations belonging to the cluster population take part in the cluster initiative? Which are not represented and why?
- What is the role of those organisations which take part in the initiative in the capacity of members or sponsors, but are neither a firm nor a research organisation?
- In which ways are the cluster initiative's member organisations committed to the initiative's proclaimed targets? Is their com-

mitment limited to the receipt of (research) subsidies? Are there cases of a purely formal membership?

- Which are the competences of the cluster initiative's organisational bodies? In which matters are the initiative's leading personalities really entitled to speak for the cluster actors?
- How is the cluster organisation controlled by the cluster board(s) (if available) and by the member organisations?
- Are there apparent differences in the distribution of benefits between cluster actors? Which of the participating organisations benefit more, which benefit less or not all?

ORGANISATION (O)

In order to be effective, cluster initiatives depend on an organisational structure which allows them to collect ideas from cluster actors, to create and strengthen team spirit, to formulate a common strategy, to manage the cluster affairs and to coordinate the joint actions of their members towards the shared target. This organisational structure comprising a more or less extended set of organisational bodies is object of the following discussion.

As cluster initiatives act within a legal organisational framework, one task of their evaluation is to assess the organisational structure and its role for the activities and results of cluster policy. Cluster initiatives are inter-organisational networks of firms and research organisations which mostly emerge and develop in the course of a self-organised process.

This process is carried by some (or in the ideal case all) cluster actors. To coordinate and steer the common activities of the initiative's members, specialised organisational bodies such as a cluster board and a cluster management organisation are created which are responsible for the planning and execution of the agreed-upon activities. The evaluation's task consists in assessing if and how the created organisational units are able to meet these requirements.

Organisational research, esp. studies on inter-organisational networks (e.g. Sydow, Windeler 2000), deliver plenty of useful hints on relevant aspects like for instance, the steering of networks, the role of facilitators and gatekeepers or the problem of organisational path dependence. Anyone who thinks, however, that organisational theory would offer an ideal solution for the organisational structure and processes of cluster initiatives will be disappointed (Kühl 2015). Organisational theory does not offer one "royal road" for the organisational construction of cluster initiatives or any other organisation. Nevertheless, imitating "best practices" which is propagated by the practice-oriented cluster literature (e.g. Scheer, von Zallinger 2007; Sölvell 2009; Lindquist, Ketels, Sölvell

2013) is a reasonable way to find pragmatic answers to organisational problems and to avoid mistakes. It produces if not optimal, at least workable solutions.

In the interaction between the business literature and practical life of cluster organisations, a general organisational structure of cluster organisations has developed (figure 2). It represents a mix of functional elements (functions realised by the cluster organisation) and organisational elements (organisational structures in which the functions are performed). In respect to functional elements, the cluster organisation has to address and adapt strategical long-term cluster management goals (strategical cluster management) and address operative activities that relate to strategy implementation. These functions are executed by different parts of the organisation (e.g. boards, a cluster management organisation or working groups). In respect to the legal organisational form, many cluster organisations in Germany chose the legal form of a registered association (eingetragener Verein), while in other cases often a private limited company (Gesellschaft mit beschränkter Haftung) is established for the cluster management organisation.

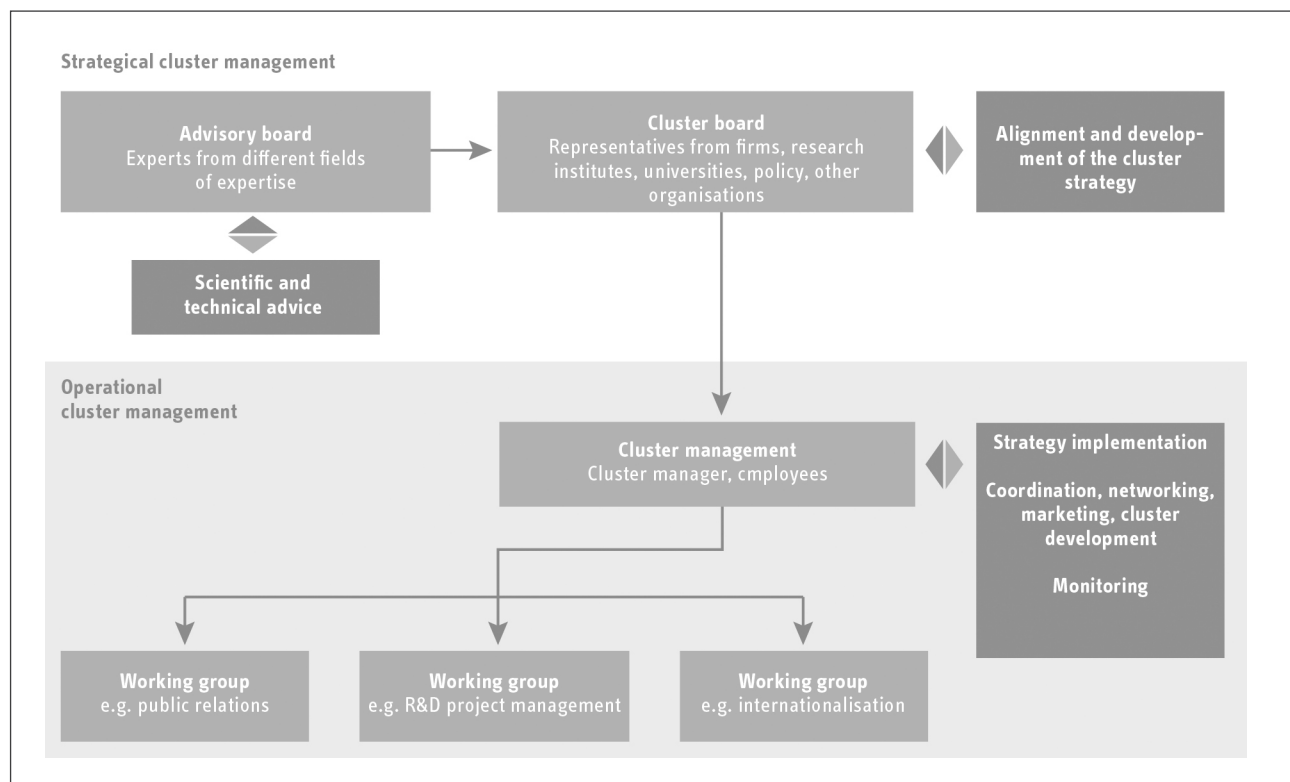


Figure 1: Generalised Graphical Representation of the Organisational Structure of the Leading Edge Clusters

Source: Rothgang et al. 2014, p. 53, own depiction based on Scheer, von Zallinger 2007, Part B, p. 59.

It is important to note that cluster organisations are not as like as two peas in a pot. In contrary, most practical cluster organisations differ from this rather general structure in different manners: Often, the organisational structure is by far not as differentiated as this structure implies (e.g. in respect to boards or working groups). There are also cases in which no formal cluster organisation exists at all. Cluster organisations are in

many instances run by some core firms and research organisations, in others they are located within a public administration (e.g. when they are developed in order to foster regional development).

In industrialized countries like Austria or Germany, cluster organisations do not develop within an empty organizational landscape. Federal

states (Bundesländer) and regional administrations promote regional development, different agencies implement public programmes or offer advice for firms. In addition, chambers of commerce as well as industrial, business and trade organisations offer services that overlap with activities of cluster organisations. The activities of research and innovation clusters within long-established industries⁵ are often related with other (frequently older) organisational structures as, for instance, the industrial collective research networks in Germany. For the evaluation this means, that the role of cluster organisations for regional development can only be understood and assessed by taking their organisational environment and the relating overlap of functions and activities into account, which mostly differ from region to region or from sector to sector.

Practical experiences collected in the evaluation of cluster policies offer some hints as to the evaluation of organisational structures:

- There is no easy way to link the formal organisational structure of a cluster organisation to performance or success of the cluster initiative. A cluster organisation can be more or less successful, independent on whether it is located in a public business development agency or not, whether the organisational structure is more top-down or not, whether the cluster management plays a core or marginal role in the total organisation.
- One important factor of success of a cluster organisation is the activity level of the organisation (process dimension). Activity level can be understood as activities of the cluster management, involvement of individual core actors and relationship between different actors. With no common activities, cluster organisations cannot be successful because no relevant results can be obtained. In contrast, a high activity level itself does not guarantee success of the cluster organisation.
- Cluster organisations change over time, regarding activity level and internal processes, partly also in respect to their formal organisation that can develop away from public administrations or in respect to cooperation with other (cluster) organisations. In order to be successful, cluster organisations do not have to live until eternity. However, if they pursue to address the needs of their members, it is important to keep the activity level high.
- Most cluster organisations that can be found resulted from or were associated with public funding programmes. At the same time, they address public tasks (like fostering regional development). When public funding expires, it is very probable that they either vanish or change their character towards organisations that usually do not serve public targets any more (like being transformed into a common research facility for the member firms).

Tasks in the evaluation of cluster initiatives are to understand both their role as intermediate organisation within a regional innovation environment and the internal processes and activities that are fostered by these organisations: Especially the following questions deserve attention:

- How do evaluators estimate the ability of the cluster organisation to act effectively (by taking into consideration own findings,

the impressions transmitted by cluster actors and, if available, the assessments of public agencies responsible for the programme's implementation)?

- How is the organisational landscape of the cluster area changed by the development (or public promotion) of a cluster organisation and what functions do these organisations fulfil?
- Is the organisational solution chosen satisfactorily suited to the individual situation of the cluster initiative (field of technology, local surroundings, core actors, relationship with policy measures and administrations)?
- What is the present and future financial basis of the cluster organisation?
- Who are the central players of the cluster initiative's organisational structure? What about the power relations between the cluster management and different (groups of) actors? Does a more democratic or a more oligarchic mode of decision-making prevail?
- In which ways do cluster member organisations take part in the activities organised by the cluster organisation and make use of services offered? How are these aspects assessed by the cluster actors?
- Which factors do determine the focus and level of the initiative's activities over time?
- Which path dependencies in the long-run evolution of the cluster initiative can be observed? Do the newly created organisational structures create the impression that they could be sustainable, respectively survive the phasing-out of the cluster programme?

FUNCTIONALITY (F)

Functionality addresses the ability of a cluster initiative to fulfil the set of tasks assigned to it by its initiators. Analysing functionality means scrutinising and monitoring (only in the case of accompanying evaluations) the cluster initiatives' activities and its demonstrated organisational and strategic capabilities. However, in this field the evaluation task is not restricted to merely counting number of activities such as meetings, seminars, advanced training courses and so on. A central part of the evaluation task consists in scrutinising the relevance of the observed activities with regard to the initiative's targets.

The first basic task of cluster initiatives operating in the context of cluster programmes consists in establishing and maintaining good working relationships with the responsible government authorities on the one hand and the cluster actors on the other hand, thus connecting donors and addressees of public funding. This task includes the obligation to mediate in conflicts between cluster actors. The second central and more difficult task consists in bundling ideas, common targets and expectations of the cluster actors, to formulate a set of shared aims and mobilise the required internal and external resources. Furthermore, the

cluster management is called on to practically support cluster actors in fulfilling the (necessary) bureaucratic obligations which are connected with the receipt of public subsidies. To make the cluster visible in external communication processes is another task of the cluster initiative managers. Apart from this, the cluster management offers a varying range of service activities, the use of which may be restricted to formal members or free to all cluster actors independently of their actual member status.

Our evaluation experience of the LECC showed that the capabilities of cluster managements to establish and maintain good working relations with the relevant project management organisations (Projektträger) and the cluster actors were differently pronounced. Relationships between the public agencies and the cluster managements were generally relaxed, but in some (rather exceptional) cases the project management organisations intervened in the cluster's affairs to bring the initiative back on course. The strategic work of cluster managements that was required to develop further the cluster strategies was of differing quality. The same applies to the external appearance of the cluster organisations. All initiatives offered a series of services for the cluster members which were partially open to non-members. The portfolio of services offered varied with the special environments of the cluster initiatives.

Evaluation experiences show that those activities which can be easily counted and used in quantitative comparisons and analyses find more attention than such aspects that are primarily qualitative in their nature. Cluster organisational bodies help the cluster actors to correctly fulfil the bureaucratic regulations in programme implementation, help them to find research partners inside and (sometimes) outside the defined cluster population, intervene in conflicts between cluster partners, organise technical seminars and continuing education events. Occasionally, they are also responsible for the maintenance of common research facilities. And they play a core role in formulating common cluster strategies, in advancing their execution and strengthening the "collective spirit" required for the sustained engagement of the cluster members (for the last Dalum et al. 2005, p. 241).

Major questions that should be analysed with regard to the operation of the cluster organisational bodies are:

- Does the cluster organisation fulfil its role in implementing the programme on the cluster's level? Are the technical operations (management of cluster projects, delivery of written reports at due time) carried out satisfactorily?
- Do the responsible cluster bodies (e.g. cluster board, strategic committees) give due attention to the strategic advancement of the cluster initiative? Which role do strategy documents (if available) play in the mid-term planning of joint projects? Do organised deliberations between cluster bodies' representatives and the whole community of cluster actors take place with regard to the long-term development of the cluster?
- Do the cluster bodies' leaders succeed in creating and strengthening the "collective spirit" of the initiative's members as well as the whole cluster population?

- How do the cluster bodies fulfil their expected role towards external addressees (public relations, network building)?
- Do the cluster organisers support cluster actors in establishing a regular exchange and flow of information between cluster projects and cluster actors who are not directly involved in the execution of the projects?
- What kind of services does the cluster management offer for the cluster actors? Do these services correspond to the specific needs of the cluster actors? Are they used by their addressees in a reasonable degree and how do they assess the quality of the services?

PERFORMANCE (P)

Performance measurement and explanation is ultimately the core task of any evaluation, whether the object of investigation is a policy measure, a regulation or an intermediate organisation. Measuring performance of cluster initiatives is important for summative evaluation (evaluating the results from the existence of the cluster initiative in order to legitimate its existence and funding), but also for formative evaluation (in order to increase the effectivity of the cluster initiative).

The "performance" of a cluster initiative may be comprehended either in a pragmatic way by designating the verifiable degree of fulfilment of its self-defined short- and medium-term aims and tasks (pragmatic perspective of the cluster initiative). Here, divergent estimations by the cluster initiative's authorities and single cluster actors are conceivable and even probable. Or, it can address the wider context of the initiative's final purpose which is defined in the framework of the cluster programme such as developing a technology cluster to one of the leading-edge clusters in the world (general impact perspective). Cluster policy evaluations will mostly aim at the second version of "performance", but must not avoid to take the first one into consideration which is far easier to handle by evaluators.

Cluster initiatives act in most cases in the context of cluster programmes and serve as preferred instrument to further the clusters' development. Thus, they are embedded in a larger setting and pursue targets that point far beyond the "well-being" of the initiative per se. Figure 3 shows the effects that are to result from the impulse of the LECC in Germany. This instrument is characterized by an open jury process in which innovation oriented cluster initiatives could apply for federal funding. The programme funded and still funds projects (mainly research projects) that are performed by the actors of the cluster initiatives. The policy impulse that results from the programme is due to an interaction of the R&D project funding and the impulse for the development of the cluster initiatives that were successful in the programme.

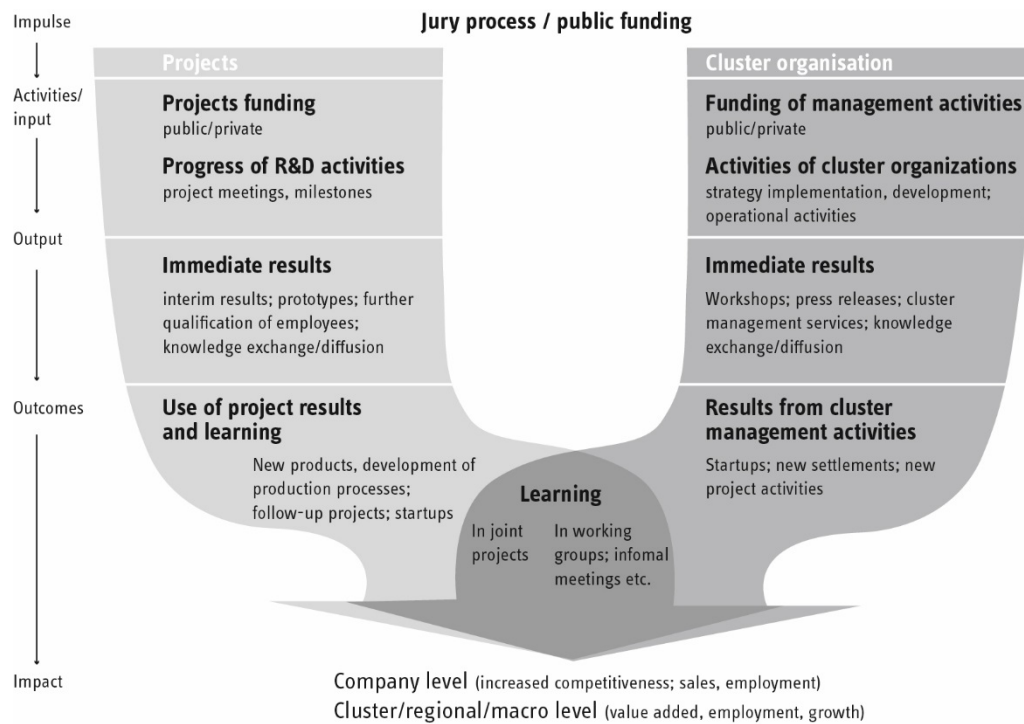


Figure 3: Effects of a Cluster Promotion Measure: The Leading-Edge Cluster Competition
Source: Own depiction.

The literature on performance evaluation of innovation policies (and other policies as well) is oriented towards answering the “counterfactual question” in order to identify the causal effects. In respect to the existence of cluster initiatives, the question would be: What would have been the development of the cluster if the cluster initiative had not come to existence? This counterfactual question could be answered by using quantitative methods. However, if the data available are not sufficient or too much intervening factors would make the quantitative analysis impossible, a qualitative assessment based on observations and propositions on probable cause-effect-relationships would be used (similar approach in Kaufman, Mayer 2011). In addition to these counterfactual analyses, also other kinds of instruments have to be applied that can be used to identify strategic targets and assess the attainment of these targets by cluster organisations. Thus, qualitative methods play a central role in the mix of methods used for the analysis and evaluation of cluster initiatives.

From the pragmatic perspective of the cluster initiative, the common set of targets and the contribution of the cluster initiative to the attainment of the cluster strategy is the core benchmark. The – mostly general and qualitative – strategy targets have to be broken down into measurable indicators. Experiences from attempts to assess the performance of cluster initiatives from the LECC show, that the initiatives especially had problems in respect to three aspects: (i) Finding meaningful indicators for showing the development of the cluster initiatives; (ii) getting a realistic view on what the initiatives could achieve, and (iii) identifying causal ef-

fects of the cluster initiatives’ existence and activities. In order to address these issues, different more or less formalized routes were taken. For example, the cluster initiative MicroTEC Südwest used a balanced scorecard in order to assess strengths and weaknesses of the cluster region in order to set targets for a future cluster development (Fachverband MST BW n.d.). Other cluster initiatives either relied on sector-specific software solutions (e.g. for biotechnology) or on self-developed indicator systems. In each case, the indicators and tools to be used had to be adapted to the individual situation of the cluster initiative.

Individual cluster actors assess the contribution of the cluster initiative in relation to their own target system. As individual firms or research organisations take part in the cluster initiative, there is in general a more or less strong overlap between their strategies and targets and the strategy of the cluster initiative. However, these are in any case not identical. Cluster actors decide on their participation and engagement in the cluster initiative on basis of the overlap in targets and benefits they expect from the activities of the initiative.

The **general impact perspective** on the cluster initiatives’ performance raises questions that go far beyond the discussed aspects. They include e.g. technology and industry evolution, market constellations, national-economic developments, political framework conditions and random factors influencing the evolution of the cluster population. Thus, it may be very difficult to verify and quantify any sustainable effects of a cluster initiative, even if it turns out to be more than a short-lived pheno-

menon. Nevertheless, while the self-assessment of cluster initiatives is mostly primarily oriented towards concepts from the business literature, the economic performance of cluster initiatives in respect to their welfare effects is often part of the evaluation of cluster programmes in which they function as intermediaries.

The following observations were made in the course of the evaluation of cluster initiatives in the evaluation of the LECC:

- Although cluster initiatives are often rather closely related to cluster programmes, the evaluation of the cluster initiative is not to be regarded identical to the evaluation of a cluster programme. The organisation embraces activities and services for the cluster members and provides a platform for different kinds of common activities (like research facilities, working groups). The economic results of these activities are often quite difficult to grasp (either by using qualitative or quantitative evaluation techniques). Cluster policy then aims at giving an impulse for the development of the cluster initiatives and promoting the development of the cluster actors.
- Immediate results of the activities of cluster organisations (new cooperations, projects or additional inputs from the cluster actors) are often rather easy to identify. This is also the case for new firms that settle in a cluster region or start-ups. However, in many cases it is difficult to identify the causality of changes in the cluster population and the activities of a cluster organisation. This problem becomes even more severe for economic impacts (development of employment or turnover). In most cases, it is simply not possible to assess the causal relationship the observed developments and the activity of a cluster organisation because many intervening factors exist. The resulting conflict in the evaluation of cluster initiatives cannot be solved: Causal relationships can only be established for a few indicators (at best). Cluster initiatives can at the same time give impulses in multiple dimensions. For these dimensions, indicator values can be collected, but evidence on causality can at best be collected by using qualitative information.
- Activities of a cluster organisation have an aspect of “public entrepreneurship”: Searching for activities and services that produce value for the cluster actors is part of that story (of course in addition to looking for public funds and programmes). Success in that pursuit is not guaranteed in advance and thus the risk of failure (or success) prevails.

These observations show that the economic evaluation of cluster initiatives raises complex issues which are not satisfactorily solved so far. The research community specialised on clusters and cluster policy could simply not yet produce hard evidence on the final success or failure of ambitious cluster programmes because of different factors like time patterns, complexity and uncertainty of results. Evidence-based evaluations could gradually contribute to fill this research gap.

SUCCESS AND FAILURE OF CLUSTER INITIATIVES: CONNECTING A, O, F AND P

Success and failure of cluster initiatives are, as we saw above, dependent on many external factors that are not connected to its characteristics. It would be illusive to believe that a high “quality” of a cluster initiative on its own could guarantee its success in furthering the cluster’s evolution. Nevertheless, organisational features of the initiative may have an enduring positive influence on the cluster development, provided that necessary structural prerequisites do exist.

Against this background, the discussed characteristics of cluster initiatives come into play. The mutual, dynamic interrelationships between agency, organisational structure, functionality and performance are a rather challenging research topic for the construction of middle-range theories on cluster initiatives/clusters and cluster policy as well as for ambitious evaluations. Theory-driven evaluation will start from a set of theoretical assumptions (working hypotheses) on these mutual relationships between A, O, F and P. This pattern corresponds to the C-M-O model (context – mechanism(s) – output) of “realist evaluation” (Pawson 2013, pp. 21-28).

The following observations were made in respect to the relationship between agency, organisation, functionality, and performance that should be kept in mind in the assessment of cluster initiatives:

- Cluster initiatives thrive on the active involvement of cluster actors and an active cluster organisation and cluster management. Thus, agency, a thriving organisational structure and adequate functionality are one precondition for a successful cluster initiative that has a positive influence on the development of the cluster actors and thus also of the cluster.
- At the same time, an active cluster initiative does not ensure a positive performance of the cluster actors and the cluster. Prevailing uncertainties in respect to success and impact of the activities that are initiated precludes such a simple relationship to evolve.

To make matters even more difficult, there are different ways how cluster initiatives can possibly be successful (or unsuccessful): Positive developments might be initiated by a close cooperation of some core actors (like large firms, research organisations or innovative small or medium sized enterprises). Alternatively, a successful cluster initiative might result from an active cluster management or a cluster manager who develops a business idea how the initiative might serve the demands of cluster firms. Alternatively, a positive development might also result from a rather successful R&D project that develops into new business, employment or value added creation strategies.

Cluster initiatives develop within a complex framework and lead to different impulses on the clusters and cluster regions. For the evaluation, this means that (i) the evaluation should be open to observe which

effects result from different dimensions that make up cluster initiatives; only focussing on one aspect (e.g. the results of common research activity) could mean that we oversee relevant aspects; (ii) different structural patterns of cluster initiatives could lead to positive results (whether more focused on core actors or a strong cluster management; whether it has developed in a cooperation of cluster firms and universities or it is originating from an attempt of a regional policy agency); (iii) because of the different structures of cluster initiatives that can be observed, there is also no uniform framework for the evaluation of cluster initiatives. The evaluation framework has to be adapted to the individual initiative to be assessed.

CONCLUSIONS

As a result from the experiences with the performance measurement of cluster initiatives, the following conclusions are drawn with regard to their evaluation:

- Understanding the logic of cluster initiatives is crucial to understanding how cluster programmes work and what their (possible) effects are. The fact that cluster initiatives in most cases experience a life outside and after the cluster programme in which they are used instrumentally, suggests that the initiatives and the related cluster promotion programme should be regarded separately.
- A low degree of representation of the cluster population by the cluster initiative might be a serious problem for many of the hundreds of cluster initiatives which have been created in the industrial countries during the last decades. Some cluster programmes such as the German LECC have avoided this problem by applying a careful screening and selection procedure.
- Cluster initiatives may adopt widely varying organisational shapes that all can be suited to the cluster's specific features. To become effective, they all need a workable organisational structure. Critically scrutinizing the workability and functionality of the created organisational bodies is a challenge for evaluators.
- Both accompanying and ex post evaluation of cluster activities have their own value in respect to policy: the economic impacts of cluster organisations and cluster policies can only be comprehensively assessed in an ex post perspective (in case of cluster policies at the earliest about three to five years after the end of funding). At the same time, many aspects that regard processes, activities, network development, can best be assessed by an ongoing evaluation.
- The question of causality should be addressed (whether the effects that are observed are in fact due to the cluster initiative's activities).
- The aspect of time lags of effects is extremely important in the evaluation of cluster programmes. Activity and input indicators can be measured rather early. However, it is not clear how they are related to policy success. Impact indicators that are relevant for policy success only change due to the policy impulse after some years have passed.
- The task of assessing the performance of a cluster initiative can be improved, if the evaluation is based on an impact model that shows the assumed timing and kinds of effects. During the evaluation, when additional information comes up, this model can be refined.
- Substantial ex ante uncertainty prevails in respect to the success of the policy and the effects that result from it. This is due to the fact that the policy impulse interrelates with different other impulses from outside the clusters (e.g. market and technological development) and at the same time meets rather different situations in the different cluster organisations and clusters (firms involved and their strategies, different abilities of cluster organisations to promote cluster targets etc).
- At the same time, measuring the effects of cluster programmes poses substantial challenges. Input and activity effects still can be measured quite easily, but their relation to policy success often is unclear. Outcome and impact effects are often difficult to relate to the policy because of many intervening variables.
- In practical evaluation, a balance has to be found between the necessary effort of a comprehensive evaluation and the means available for the evaluation task. It is important that this decision is made under consideration of the complexity of the evaluation challenge.

This leaves the evaluation of cluster programmes with the challenge not only to understand the contribution of cluster initiatives that are used as intermediaries but also the interdependence between both of them. And, of course, the final measure of policy success is not whether the programme at hand leads to an activation of cluster actors to participate in the cluster initiative, but the final outcome in respect to economic impact that evolves in the medium- and long-term.

REFERENCES

- Andersson, T., Schwag Serger, S., Sörvik, J. and Wise Hansson, E.** (2004), *The Cluster Policies Whitebook*, IKED – International Organisation for Knowledge Economy and Enterprise Development, VINNOVA, The Competitiveness Institute, printed by Holmbergs.
- Dalum, B., Pedersen, C.Ö.R. and Villumsen, G.** (2005), *Technological Life-Cycles: Lessons from a Cluster Facing Disruption*. *European Urban and Regional Studies* 2005(12): 229-246.
- European Commission, Enterprise and Industry** (2012), *Global Cluster Initiative Survey 2012. Survey Summary Report*, September 2012, Brussels: European Commission.
- Fachverband MST BW** (n.D.), *MicroTEC Südwest. The Leading Cluster For Smart Solutions. Roadmap 2020+*.

Internet Source. Url: http://microtec-suedwest.de/fileadmin/Roadmap/140428_Roadmap.pdf, accessed at 19/11/2015.

Kaufmann, P. and Mayer, S. (2011), Is there a use for counterfactual impact evaluation? In: *fteval Journal for Research and Technology Policy Evaluation*, issue 37, December 2011, pp. 29-43.

Kühl, S. (2015), *Sisyphos im Management. Die vergebliche Suche nach der optimalen Organisationsstruktur. 2., aktualisierte Auflage.* Frankfurt/New York: Campus Verlag.

Lindquist, G., Ketels, C. and Sölvell, Ö. (2013), *The Cluster Initiative Greenbook 2.0.* Stockholm: Ivory Tower Publications.

Pawson, R. (2013), *The Science of Evaluation. A Realist Manifesto*, London, Thousand Oaks, CA, New Delhi, Singapore: Sage.

Porter, M. E. (1998), 'Clusters and Competition: New Agenda for Companies, Governments, and Institutions', in M.E. Porter, *On Competition*. A Harvard Business Review Book. Cambridge, MA, London, England: Harvard University Press, pp. 197-271.

Rothgang, M., Cantner, U., Dehio, J., Engel, D., Fertig, M., Graf, H., Hinzmann, S., Linshalm, E., Ploder, M., Scholz, A.-M., & Töpfer, S. (2014). Begleitende Evaluierung des Förderinstruments "Spitzencluster-Wettbewerb" des BMBF, Abschlussbericht. [RWI - Rheinisch-Westfälisches Institut für Wirtschaftsforschung, ISG – Institut für Sozialforschung und Gesellschaftspolitik GmbH, JOANNEUM RESEARCH Forschungsgesellschaft mbH, Graz, Lehrstuhl für Volkswirtschaftslehre/Mikroökonomik der Friedrich-Schiller-Universität Jena im Auftrag des BMBF]. Essen: RWI.

Rothgang, M., Cantner, U., Dehio, J., Engel, D., Fertig, M., Graf, H., Hinzmann, S., Linshalm, E., Ploder, M., Scholz, A.-M., & Töpfer, S. (2015). Accompanying Evaluation of the Funding Instrument "Spitzencluster-Wettbewerb" (Leading-Edge Cluster Competition) of the Federal Ministry of Education and Research. Final Report - Summary. RWI Materialien 90. Essen: RWI.

Scheer, G. & von Zallinger, L. (2007), *Cluster management. A practical guide.* Eschborn: GTZ.

Sölvell, Ö. [2008] (2009), *Clusters. Balancing Evolutionary and Constructive Forces.* Second Edition. Stockholm: Ivory Tower Publications.

Sydow, J. (2003), Dynamik von Netzwerkorganisationen – Entwicklung, Evolution, Strukturierung. In: Hoffmann, W.H. (Hrsg.): *Die Gestaltung der Organisationsdynamik.* Stuttgart: Schäffer-Poeschel, pp. 327-356.

Sydow, J. and Windeler, A. (Eds.) (2000), *Steuerung von Netzwerken. Konzepte und Praktiken.* Opladen/Wiesbaden: Westdeutscher Verlag.

Voelzkow, H. (2014), Regionale Wirtschaftskluster zwischen dem deutschen Produktionsregime und der Governance von Wirtschaftssektoren: Stand der Forschung zu den 'Varianten des Kapitalismus' und einige Schlussfolgerungen für die Clusterpolitik. In: Beck, R., Heinze, R.G. und Schmid, J. (Eds.), *Zukunft der Wirtschaftsförderung.* Reihe Wirtschafts-

und Sozialpolitik 14. Baden-Baden: Nomos-Verlagsgesellschaft, pp. 135-165.

AUTHORS

DR. MICHAEL ROTHGANG

RWI - Rheinisch-Westfälisches Institut für Wirtschaftsforschung, Essen

E: michael.rothgang@rwi-essen.de

DR. BERNHARD LAGEMAN

E: bernhard.lageman@t-online.de

AGENCIFICATION REVISITED A CASE FOR THE EVALUATION OF FFG AS AN AGENCY

RUPERT PICHLER

INTRODUCTION: AGENCIES AS A NPM TOOL

Agencies are an essential element out of the New Public Management (NPM) tool box. Funding agencies also count among the most important intermediaries in the field of research policy. When the NPM wave began to roll in the 1990s agencification was a centrepiece there as well. In the NPM context agencies are deemed to ensure better goal attainment through greater efficiency. Basically, the latter is achieved by the division of labour and clarification of roles as according to the NPM logic agencification means separating regulatory from operational tasks (Christensen and Lægheid, 2006; Talbot, 2004; Bach und Huber, 2012).

One underlying model is the principal-agent approach: the principal delegates a task to the agent who knows more than the principal about a specific field where the principal wishes to intervene (Braun and Guston, 2003). Such a relationship is typical for the implementation of research policies which *“needs the co-operation of scientists to overcome the implicit lack of knowledge or, in other terms, the ‘information asymmetry’ inherent in functional differentiation.”* (Braun, 2003, 310).

When the Austrian Research Funding [official translation: Promotion] Agency FFG was created in 2004 this was the result of a reform that was much about NPM-style clarification of roles and the reshaping and merging of previously existing funding institutions. Therefore, the foundation of FFG was widely conceived as agencification. Assuming that the main goal of FFG was providing a more efficient structure being more responsive towards the government (where existing RFOs were to be integrated) as well as the applicants (where ministerial funding was to be outsourced), possible evaluation questions may be fairly simple: Did FFG live up to the expectations at the time of its foundation? Is FFG in an appropriate position in order to fulfil today’s mission? Does it serve the applicants in a way suitable for this mission? Does it have appropriate managing and governance structures for that purpose?

These questions may seem trivial. However, most of the recent evaluations addressed the impact level of research policy measures including

“customer” satisfaction. This is in line with the common NPM narrative according to which a higher impact will be achieved through NPM. Nonetheless, NPM approaches like agencification often mean a massive change of administrative structures. Therefore, a shift of the evaluation focus to functional efficiency of both the principal–agent play and the internal set-up of FFG – as reflected in the questions above – might be justified in order to determine the cost of what may indeed be greater effectiveness of agencies. This is why we should look at FFG specifically as an act of agencification rather than just some form of research funding organisation (RFO). Since its foundation in 2004 FFG has never been subject to such an evaluation. This article tries to outline a framework of what we already know and where a possible evaluation should be focused upon in terms of agencification.

1. AGENCIFICATION IN RESEARCH FUNDING: A TYPOLOGY

When we set out to do so, we obviously need criteria to define an agency so as to distinguish it from other intermediary RFOs. Mature innovation systems usually dispose of such intermediary bodies whose key purpose is to connect with the stakeholders rather than the division of labour as in the agencification logic.¹ This dichotomy needs to be addressed in evaluation studies if we want to learn about agencies’ efficiency in comparison.

Generally, an RFO can be defined *“as a non-ministerial institutional structure in the public domain designed to manage the allocation of government funds to research projects, individuals and/or institutions. Its governance structures enjoy considerable autonomy from the government and integrate those addressed (e.g. the scientific community) into its decision-making”* (Stampfer et al., 2010, 766). The concept of agencification fits into this definition, but it is not the whole story, not least because most RFOs have been established long before NPM was introduced.

Beyond the reasons *why* the implementation of certain tasks is separated from ministries a differentiation can be made in functional terms

¹ According to Braun and Guston (2003, 303) „[f]unding agencies were, since their origins, designed to work out and implement research policies, in preference to the usual public bureaucracy that lacked the necessary direct

along the modes of delegation – delegation by trust versus delegation by contract (Braun, 1997 and 2003). Two basic categories of RFOs can be derived from that (Stampfer et al., 2010, 767):

- The council, which manages its funds autonomously and is run by members of the communities that it addresses.
- The agency, which provides the government with managerial expertise and capacities and manages its funds on the government's behalf.

Considering these categories in case FFG is evaluated as an agency, criteria are needed against which the respective functional dimensions can be evaluated in comparison to council-type RFOs. The figure below shows how functional differences may be expressed for criteria defining intermediary functions.

intermediary function	autonomous council	agency
legitimation	corporative	technocratic
relation to the government	trustee	service provider
role of those addressed	members	customers
role of governing bodies	interest grouping	provider of expertise
expected benefit	shared responsibility	division of labour

Comparison of intermediary functions

Source: Pichler (2014) based on Braun (1997).

It becomes fairly clear that the roles according to these two models can be well differentiated. In this sense it is incorrect that in Austria the RFOs are commonly referred to as “the agencies”. It has also to be mentioned that historically they often represent subsequent stages of institutionalising research funding. Looking at FFG, we will see the relevance of this distinction in its historical dimension.

2. AN EVALUATION RESULT: “REVERSE” AGENCIFICATION AS ONE RATIONALE FOR FFG

Remarkably, the foundation of FFG was strongly associated with an evaluation exercise. While the overall story has been sufficiently discussed elsewhere (Stampfer et al., 2010) we shall focus here on how the issue of agencification was dealt with. Subjects of this evaluation in 2004 were two RFOs – the Science Fund (FWF) and the Industrial Research Promotion Fund (FFF). Both represented the “council type” RFO, nothing unusual in the case of FWF. FFF was different as its autonomous status was internationally unusual for the type of research it funded.

The starting point of the debate leading to the evaluation was the fragmentation and multiplicity of the funding system. But soon the autonomy of FWF and FFF got into the focus as an element of petrification

and lacking accountability: it seemed obvious that “most of the innovation in Austrian innovation funding system has taken place outside FFF” (Arnold et al., 2004, 57). The reason for that was that the funds “were given a strongly ‘autonomous’ status, which is to say that they were given governance structures that were dominated by their beneficiaries, rather than by ministries. This reflected a lack of confidence that they could be kept free from inappropriate, detail-level interference ... What the Research Promotion Act set out to do through autonomy, is elsewhere being achieved through increased clarity about the respective roles of ministries and agencies” (Arnold et al., 2004, ii). Eventually, this “proved to be a tragic flaw, because to a considerable extent it locked the Funds into their 1967 roles” (Arnold et al., 2004, 102).

Derived from this analysis the evaluation report recommended “the balance of power in its governance structure should be changed so that more representatives of the taxpayers and not the beneficiaries have control” (Arnold et al., 2004, 59, on FFF). As a result, the funds “need to be changed from quasi-autonomous bodies to agencies of the government” where “FFF should be merged into a broader innovation agency” (Arnold et al., 2004, 113).

Yet by 2004 the structure that was now chosen for FFG had already been operated successfully by one of FFG's predecessors (TIG), based on a more timely delegation-by-contract approach. Other attempts, though, to install elements of agencification within FFF and FWF were doomed. When in the late 1990s many new programmes had to be implemented the ministry needed the operational expertise of FFF where TIG could not be used because of its specific focus. Therefore, the legal provisions for FFF were amended by a regulation adding the implementation of contracted programmes to its scope (BGBl. I Nr. 79/1999). However, the governance structures of FFF remained unchanged, thus keeping their responsiveness towards ministerial intentions at a low level. The new task just seemed to be a nice add-on, leaving the general approach of FFF unchanged. FWF got into a similar situation, but as the programmes in question had been born out of FWF's own community a legal amendment was not deemed necessary for the time being.²

Within this context of gradual, yet uncoordinated institutional change – supplemented by the introduction of the council for research and technology development as a new advisory body to the government – the evaluation of the funds was among the triggers of efforts to change the legal set-up of research funding substantially. The debate soon concentrated on what was to become FFG. The evaluation advised to leave FWF out of a possible merger while the Court of Auditors – just having examined the funds too – said the opposite (Rechnungshof, 2003), and the strong stakeholder-driven tradition of FWF as a council-type RFO delayed any agency-like change of structures.

To sum up: even if not explicitly mentioned, the idea of agencification was relevant within the context of the evaluation of FFF and FWF as well as for the creation of FFG. But the major point was what we call “reverse” agencification. Despite a certain and sometimes dodgy tradition of ministerial in-house funding its outsourcing did not cause much debate at that time anymore, not least because of the ministries' staff restrictions. While this process – entailing e.g. the foundation of TIG 1998 – was well

underway the main problem that had been addressed by the evaluation was too much rather than too little independence of research funding. Therefore, agencification proved to be a reference for reducing instead of increasing autonomy.

3. THE INSTITUTIONAL SET-UP OF FFG: PARALLEL WORLDS SURVIVING

Taking the recommendations of the 2004 evaluation further, the first obvious question to ask is whether or not these recommendations were implemented. Firstly, the research funding reform bill of 2004 (Forschungsförderungs-Strukturreformgesetz, BGBl. I Nr. 73/2004) proposed a merger of FFF with BIT, TIG and ASA. The new agency (FFG) took the form of a state-owned limited liability company. Contrary to FFF, however, the management of FFG is thus bound by instructions of the minister(s) in charge. The explanatory note accompanying the government's proposal (Nationalrat, 510 d.B. XII.GP) features many NPM buzzwords: it called for modern governance structures, a clear attribution of responsibilities, harmonisation of fragmented structures, less transaction between ministries and intermediary bodies, all leading to simpler structures creating synergy effects and more transparency.

The choice of a limited liability company is standard procedure in Austria as no specific legal form for agencies – like e.g. in Scandinavian countries – exists (Pichler/Strimitzer, 2014). The consequences of the principals' ownership in terms of their influence, however, were intended. In that sense, the government did what the evaluation had called for, including respective regulations about the balance of power in the governance structures. The board of FFG has now a majority of members appointed by the ministries (excluding the employees' representatives), compared to the non-voting position of their (few) representatives in FFF times. In the same vein, the proposed merger of FFF into a broader agency was followed up, explicitly justified by the expected gain in efficiency during the legislative procedure.

But is this the whole story? While we will cast a glance at FWF below, there is more than one story line for the case of FFG. Agencification is about delegation of power. Even if the most important aspects of the foundation of FFG were the "reverse" agencification of FFF and the merger of existing intermediaries rather than outsourcing of ministerial programmes, the point still is that FFG represents an intermediary organisation at arm's length from the government. The tasks FFG is assigned with are therefore delegated from the government, whatever the previous situation. A look at these tasks reveals further elements that could be evaluated against the criteria of agencification.

As a funding agency, the principal task of FFG is to implement the spending of those funds. However, what the 2004 law regulated in the first place was the delegation of management responsibilities. It says little about the delegation of financial responsibilities. The level of management delegation results from the principles guiding a limited liability company, as described above. As regards the control over the allocation of funds things are less clear. The law only suggests that FFG

is to allocate funding on its own as well as on the government's behalf. In legal terms, this principle is only fully understood together with other legal provisions governing research funding (e.g. Forschungs- und Technologieförderungsgesetz). Hence, different modes from the past are surviving: FFF was autonomous not only in management, but also financial terms; both TIG and ASA were limited liability companies, but TIG got injections of capital while ASA handled the money only on behalf of the government (like FFF when it was also assigned to manage additional programmes).

FFG inherited parts of these different approaches. FFG disposes of funds underlying its own allocation decision as well as funds where it only executes the respective ministry's allocation decision. This puts the principals in a dual role as both owner and client. It remains to be determined by a future evaluation how this affects the functional efficiency of FFG. At any rate, this dichotomy feeds back to the delegation of management tasks as the different levels of financial responsibility require different ways of planning and reporting.

Summing up again, we have found some evidence of incomplete agencification: while the autonomy of FFF was reversed to the level of a state-owned company, FFG was not empowered to manage all the other programmes and respective funds on that very level. Although in-house ministerial programme implementation has almost ceased to exist, the overall judgement of the agencification success through FFG is therefore not yet spoken: did the apparent shortcomings in fact hamper FFG to the disadvantage of its function, contrary to the recommendations of the 2004 evaluation?

4. A LOST DECENNIUM? AGENCIFICATION IN POST-2004 EVALUATIONS

Neither the law establishing FFG nor any other legal provision at that time required that FFG be evaluated after a certain period of time. Most evaluations carried out in these years largely focused on programmes and their impact. The institutional perspective of research funding was featured again in the system evaluation of 2009 (Aiginger et al., on ministry-agency governance Mayer et al.). Even though this evaluation primarily addressed the effect of the research funding system on the target groups' behaviour, the underlying structures were also looked at.

Quite in line with our hypothesis the evaluation stated: *"Agencification has been started but has not been fully implemented"* (Aiginger et al., 2009, 140). This seems to have been judged as a flaw compromising the division of labour between ministries and FFG, amongst others *"caused by the existence of two parallel steering mechanisms"* (Aiginger et al., 2009, 86). After all, the owner-client dichotomy was not lost on the system evaluation. From its external perspective it was reasonable to propose that the principal's influence on the agency should generally be exerted through its right of ownership and a performance agreement. Through *"a process of earned autonomy"* this would lead *"to more operational freedom of the agencies and more room for strategic governance at the side of the principal"* (Aiginger et al., 2009, 141). However, *"that*

should be based on a prior normative discussion about the positions on and understanding of the division of labour, governance and STI policies as such" (Aiginger et al., 2009, 88).

So the system evaluation had raised questions which have also been identified above. It explicitly referred to the concept of agencification, though without discussing its appropriateness. The main concern was the ability of FFG to ensure a demand-driven "offer" of funding instruments and the corresponding paragraphs were only four of 61 in the recommendation summary part.

The issue of what happened after the structural reform of 2004 surfaced again in terms of evaluation only recently, namely in connection with FWF and the new impact assessment tool. The latter was introduced in 2013 in the course of the budget reform as a mandatory element to be applied whenever a regulation is proposed. After 2004 the traditional decision-making structure of FWF had essentially remained unaltered (except for the introduction of a supervisory board), meaning that the financial responsibilities had stayed with the president whose principal legitimation, however, is based on academic merit. In recent years this lead FWF into trouble with the new public corporate governance and budgetary requirements as well as with the court of auditors.

As a consequence, an amendment of the legal provisions for FWF was proposed. What the corresponding impact assessment (Nationalrat, 691 d.B. XXV.GP) states can be – once again – interpreted as reverse agencification, at least in part. Firstly, it was made clear that in legal terms FWF is the result of outsourcing and thus (almost) an agency. Taking into account the rulings of the constitutional court on outsourcing of public tasks the government must therefore regain more influence on the governance of FWF. Secondly, the division of responsibilities between president and management was to be clarified, and the competences of the supervisory board were to be extended in the way applying to limited liability companies. As a result, FWF got governance structures much closer to an agency (BGBl. I 110/2015). It now practically has a managing board including a CFO on equal terms with the scientific president, reporting to the supervisory board. The president and the CFO are appointed by the board, the former based on proposal by the assembly of delegates, instead of the other way round before. This fulfils the constitutional requirements as the board includes members appointed by the government. In short, the old council days are, to a certain extent, gone.

It cannot be said that evaluations in the past ten years would have entirely neglected institutional matters. But striving for the big picture of effectiveness seemingly did not allow for much more than side glances at the efficiency of institutions. The agencification issue and its functional dimension is not dead, though, as the court of auditors demonstrated with the report on the controlling systems of FFG and FWF (Rechnungshof, 2015): where accountability meets autonomy clarity on responsibilities and the return for the government beats the old glory of scientific self-organisation.

CONCLUSION

For almost two decades Austria has witnessed changes in its research funding setting that can be characterised as – or were decidedly meant to be – agencification. Therefore it is appropriate to evaluate this very dimension against criteria that define it. Yet this is not about the articles of faith on NPM or the autonomy of science, a contrast often not so sharp in international comparison (Groß et al., 2010). What is simply needed are answers whether or not the design principles applied to FFG (and, more recently, in part to FWF) and reflected in some evaluation exercises live up to their promises: if what appears to have been by and large a satisfactory performance in terms of impact (as proved by numerous evaluations) has been achieved also in an efficient way.

One key promise pertinent to NPM-style agencification is clarity and simplicity to allow for better steering. In other words, complexity counts as a functional problem as such. As regards the number of agencies, FFG reduced complexity. Although one might argue that smaller agencies could be steered better and that a certain diversity may even prove advantageous³, the crucial point in 2004 was the *"increased clarity about the respective roles of ministries and agencies"* (Arnold et al., 2004, ii).

It remains to be determined to which extent this has in fact been achieved. The persistence of different delegation levels hint at some problems with the separation of strategic and operational tasks. Hence, the checks and balances between agent and principal may be jeopardised. The balance *"is more likely to be stable ... when both need the other"* (Pollitt et al., 2005: 23). Since this mutual dependence can turn out differently according to the respective delegation level, as a result the principal(s) may alternate between too much and too little guidance of the agency.

However, while undue influence of the beneficiaries on "their" agency is not a problem anymore neither with FFG⁴ nor – for the better part – with FWF, the risk of "reverse" agency capture towards the ministries cannot be fully excluded. One possible reason in the case of FFG is its nature as a multi-principal agency, potentially deepening the information asymmetry between FFG and its principals because FFG is put in a position where it has a better overview and thus more information than individual ministries.

Even if this yields positive coordination effects (Pichler, 2014) it comes at a price. Limited coordination capacities at government level have allowed FFG to fill this strategic gap without being able to entirely bridge it as an agency (as a general notion expressed by Braun, 2008, 298). The consequences are – contrary to what was intended – blurred boundaries between the principals' and the agency's tasks. Together with the coexistence of different delegation levels the options to fully exploit the potential of FFG as an agency seem to be restricted. For a comprehensive assessment of the situation an evaluation is needed. It should focus on the overall governance framework and internal core processes, and what relevance they have for the efficiency of FFG. The findings of such an eva-

3 As often argued in the case of the German „Projekträger“

4 That FFG terms its applicants „customers“ seems to reflect NPM wisdom rather than an imminent danger of agency capture..

uation are expected to supply input to the design of a coherent steering logic for FFG. Similarly, an evaluation of another big funding agency born out of a merger, AWS, might be feasible.

REFERENCES

Aiginger, K., Falk, R., Reinstaller, A. (2009). Reaching out to the Future Needs Radical Change. Towards a New Policy for Innovation, Science and Technology in Austria: Synthesis Report. Wien: WIFO.

Arnold, E. et al. (2004). Evaluation of the Austrian Industrial Research Promotion Fund (FFF) and the Austrian Science Fund (FWF). Synthesis Report. Brighton/Wien: Technopolis.

Bach, T., Huber, E. (2012). Comparing agencification in continental countries. Government Agencies: Practices and Lessons from 30 Countries, eds. Verhoest, K., van Thiel, S. Bouckaert, G., Lægreid, P., Basings- toke: Palgrave Macmillan, 203–210.

BGBI. Bundesgesetzblatt für die Republik Österreich. <http://www.ris.bka.gv.at/Bund/>

Braun, D. (1997). Die politische Steuerung der Wissenschaft. Ein Beitrag zum „kooperativen Staat“. Frankfurt a.M./New York: Campus.

Braun, D. (2003). Lasting tensions in research policy-making – a delegation problem. *Science and Public Policy* 30(5), 309–321.

Braun, D. (2008). Lessons on the political coordination of knowledge and innovation policies. *Science and Public Policy* 35(4), 289–298.

Braun, D., Guston, D.H. (2003). Principal-agent theory and research policy: an introduction. *Science and Public Policy* 30(5), 302–308.

Christensen, T., Laegreid, P. (2006). Agencification and Regulatory Reforms. *Autonomy and Regulation: Coping with Agencies in the Modern State*, eds. Christensen, T., Lægreid, P., 8–52, Cheltenham: Edward Elgar.

Groß, T., Karaalp, R.N., Wilden, A. (2010). Regelungsstrukturen der Forschungsförderung. Staatliche Projektfinanzierung mittels Peer-Review in Deutschland, Frankreich und der EU, Baden-Baden: Nomos.

Mayer, S. et al. (2009). Systemevaluierung der österreichischen Forschungsförderung und -finanzierung, Teilbericht 3: Governance in der FTI-Politik im Wechselspiel zwischen Ministerien und Agenturen. Wien: WIFO.

Nationalrat, 510 d.B. XII.GP. http://www.parlament.gv.at/PAKT/VHG/XXII/I/I_00510/fname_021201.pdf

Nationalrat, 691 d.B. XXV.GP. http://www.parlament.gv.at/PAKT/VHG/XXV/I/I_00691/fname_423860.pdf

Pichler, R. (2014). Delegierte Koordination. Zur Koordinationsfunktion intermediärer Organisationen in der österreichischen Forschungsförderung. *Österreichische Zeitschrift für Politikwissenschaft* 43(4), 329–346.

Pichler, R., Strimitzer, E. (2013). Ausgliederungen als Instrument eines modernen Verwaltungsmanagements: Grundlagen, Probleme, Praxis. *Handbuch des modernen Verwaltungsmanagements*, eds. Strimitzer, E., Thewanger, E., Bauer, J., Nayer, M., Wien: Linde Verlag, 99–135.

Rechnungshof (2003). Reihe Bund 2003/2, Wien: Rechnungshof. http://www.rechnungshof.gv.at/fileadmin/downloads/2003/berichte/berichte_bund/Bund_2003_2.pdf

Rechnungshof (2015). Reihe Bund 2015/7, Wien: Rechnungshof. <http://www.rechnungshof.gv.at/berichte/ansicht/detail/ffg-und-fwf-interne-kontrollsysteme.html>

Stampfer, M., Pichler, R., Hofer, R. (2010). The Making of Research Funding in Austria. *Transition Politics and Institutional Development, 1945–2005. Science and Public Policy*, 37(10), 765–780.

Talbot, C. (2004). The Agency idea: sometimes old, sometimes new, sometimes borrowed, sometimes untrue. *Unbundled Government: A critical analysis of the global trend to agencies, quangos and contractualisation*, eds. Pollitt, C., Talbot, C., New York: Routledge, 3–21.

AUTHOR

RUPERT PIRCHLER

Federal Ministry of Transport, Innovation and Technology

E: rupert.pichler@bmvit.gv.at

A REGIONAL FUNDING ORGANIZATION FOR BASIC RESEARCH? UNCONVENTIONAL MISSIONS AND THEIR EVALUATION

GRIT LAUDEL AND MICHAEL STRASSNIG

INTRODUCTION

In many OECD countries, the state funding for research has been transformed into a split funding mode that combines block funding for universities with competitive project grant funding. Consequently, research funding organisations (RFOs) have become obligatory points of passage for conducting research in resource-intensive fields. Grants from RFOs has also become embedded in institutionalised evaluations and have become a pivotal benchmark for the reputation of both research organisations and researchers.

Thus, RFOs emerged as central actors in innovation systems. In science studies, they are commonly conceptualised as intermediary organisations because they mediate between science policy (representing state and public interests), scientific communities and their priorities and standards, and researchers who apply for and receive grants in order to further develop their individual research agendas. RFOs are usually jointly controlled by the state and scientific communities (Braun 1993, 1998).

RFOs enjoyed significant attention of science policy and sociology of science studies which, however, focused their attention on state-level RFOs. It has been shown that the state depends on RFOs and other advisory organisations with intermediary functions to be able to govern science (Van den Daele et al. 1977, Guston 1996, Van der Meulen and Rip 1998, van der Meulen 2003). Principal-agent theory has been suggested as a theory that can explain structures of and processes in and around research councils (Braun 1993; Rip 1994; Guston 1996; Braun 1998; van der Meulen 1998; Braun and Guston 2003; Caswill 2003), but has been criticised for not being able to do justice to the complex embeddedness of research councils (Morris 2003; Shove 2003). The actual impact of RFOs on research was investigated very selectively, namely with regard to research quality (do RFOs select the best applicants?) and public policy goals (does RFO funding support the development of particular fields?), see Gläser and Laudel 2016.

RFOs that have a regional focus and thus are intermediaries between regional interests and scientific communities are less common. Examples for such regionally oriented RFOs are the Mercator-Foundation for the Ruhr area in Germany, the Einstein Foundation in Berlin, the Vienna Science and Technology Fund in Austria. The role of these regional RFOs as intermediary organisations has hardly been investigated. This neglect is due to a bias of science and innovation policy studies, which commonly

associate regional actors with interests in the application of research findings and technology development. A commonly accepted division of labour in the innovation ecosystem is that the nation state takes care of science funding while regional innovation policies are predominantly concerned with fostering transfer and innovation in economic contexts in those thematic fields that are considered of regional public interest. They address universities and other research institutions through their “third mission”, thereby seeking to push activities of research institutions “downstream” in the innovation chain and systematically neglecting the merits of “upstream activities” such as curiosity-driven fundamental research. Science policy studies mirror this interest by focusing on governance arrangements that support this aim (e.g. Feldman and Audretsch 1999, Rip 2002, Trippl and Maier 2010, Van Geenhuizen and Nijkamp 2012, Gertner and Bossink 2015).

At the same time, in many regions in industrialized countries basic research landscapes have emerged, i.e. concentrations of basic research that show a distinct profile, coherence, and synergies from local collaborations. Locality does indeed matter in that it enables collaborations depending on face-to-face contacts. After some initial attempts (Inhaber 1974), little attention has been paid to the benefits of co-location for fundamental research beyond some descriptive bibliometric studies (e.g. Hoekman et al. 2010). A bibliometric study of biotech research at the city level showed that new scientific topics in this field emerge where related topics already exist and others were more likely to disappear when they were weakly related to the scientific profile of the city (Boschma et al. 2014), indicating that a city's research amounts to more than simply the sum of its parts.

These two developments create a policy question and a corresponding research gap that can be described by the same question: How can a regional basic research landscape be developed? A more specific question, which is addressed by this paper, asks about the role RFOs can play in developing regional basic research landscapes. With this paper, we aim to contribute to closing this gap by answering the question how (by what mechanisms, with what effects) a small regionally oriented RFO can contribute to the development of a dense regional basic research landscape. We discuss the case of the Vienna Science and Technology Fund (Wiener Wissenschafts-, Forschungs- und Technologiefonds, WWTF). The paper is based on an evaluation of WWTF in 2013/2014, in which both authors were involved: Grit Laudel was part of the external expert group that evaluated the WWTF (Laudel 2013, Costa et al. 2014). Michael Strassnig is a Programme Manager of the WWTF and in that

role participated in the preparation and organisation of the evaluation (WWTF Office 2014).

The WWTF was established as a private non-profit fund in Vienna, whose mission is to strengthen the Vienna research landscape by contributing to the accumulation of critical mass of excellent scientific research in selected fields. It has been operating since 2002, and first effects of its work became visible after ten years.

Evaluating such a specific mission required a targeted micro-level approach. Comparative case studies were conducted in three selected research areas of WWTF funding activities, namely Life Sciences, Information and Communication Technology, and Interdisciplinary Mathematics. The case studies were based on semi-structured interviews with 25 grantees from all three major funding programmes who received WWTF funding. The interviews were prepared by bibliometric structural analysis of their publication oeuvre (Gläser and Laudel 2015; for more details about the methodology and methods see Laudel 2013).

1. AN INTERMEDIARY ORGANIZATION FOR A CITY'S RESEARCH

Vienna is a large research and innovation area with a total of 2.8 billion Euro annual spending on R&D of which 740 million Euro are spent for basic research.¹ It is by far the most important Austrian research landscape with nine public universities, four larger universities of applied science and almost all Austrian extra-university public research institutes. About 30.000 researchers and 180.000 students work and study in Vienna.

The WWTF was established in 2002 as a local RFO by the City of Vienna. As an intermediary organization it mediates between regional political interests, the local scientific communities and international scientific communities. This is reflected in its decision-making body (Board of Directors) and advisory board. The former consists of two WWTF officials, two representatives of the private banking foundation that funds WWTF, and two academics. It set the mission to "strengthen Vienna as a location of excellent research through funding projects and persons (WWTF Office 2014: 9). Two thirds of advisory board members are academics working in Vienna, the remaining third consists of municipal politicians (WWTF Office 2014: 10-11).

The mission of the WWTF is to promote research in Vienna by contributing to the accumulation of critical mass of excellent research in selected fields (WWTF 2008). To fulfil this mission, the WWTF has an annual budget between seven and thirteen million Euros. In addition, the WWTF manages funding programmes directly financed by the City of Vienna, namely the *Vienna Research Group for Young Investigators Programme*, the *Social Sciences and Humanities in Vienna Programme*, and a funding programme for university infrastructure.

As an intermediary organisation the WWTF has an unusual degree of autonomy. Although founded by the City of Vienna, it is not an agency of the municipal administration but a private non-profit organisation. Municipal policy makers do not interfere with the decisions of the WWTF beyond their role as members of the advisory board.

Funding programmes are designed independently by the WWTF, which constantly develops, adopts, and scrutinises ideas for new thematic areas. In its initial screening of ideas, WWTF applies two basic criteria. The proposed thematic area must be large enough to enable competitive calls, which means that it must overlap with research interests in a larger number of Viennese research organisations. Second, the scientific quality of Viennese research in this area must be well-above average. If these conditions are met, WWTF explores the specificities of the new field and possible impacts that can be achieved with different interventions by identifying and interviewing central actors including researchers and decision makers in relevant organisations. These findings from interviews are triangulated with a bibliometric analysis, interviews with experts from abroad (who are familiar with the research in Vienna but have no personal interests there) and data about locations abroad that are comparable in terms of quality, scope and scale.

The peer review of grant proposals involves the international rather than local scientific community in order to prioritize research quality and to exclude particularistic local interests (WWTF Office 2014: 20). The process is managed by the WWTF office who organizes written reviews and local meetings of international panels.

In these procedures, the local scientific and political interests are present in the WWTF's mission and boards, while the scientific profile and design of interventions is under the authority of the WWTF, and funding decisions under the authority of the international scientific community. Under these conditions, the strategy of an RFO becomes the decisive factor for its contributions to the dynamics of a research landscape.

Small RFOs operate in a complex research funding ecosystem that includes large and comprehensive RFOs. They have the disadvantage of needing to achieve effects with relatively small budgets. At the same time, they have the advantage of being able to design very specific interventions. Large publicly financed RFOs must balance a huge variety of interests involved in a national innovation system. These include the balance of targeted and researcher-driven funding, funding of a variety of collaborative forms of research, funding basic and application-oriented research, considering regional proportions and keeping national career patterns healthy. Smaller RFOs, on the other hand, have more freedom to set their agenda, particularly when they are as independent from political interests as the WWTF currently is.

The main strategy of smaller RFOs often consists in finding gaps in the funding portfolios of larger players and funding topics and/ or research with particular properties that would not find funding otherwise. Finding gaps has little to do with finding niches because the idea of niches implies protected but unimportant areas. However, since even large

RFOs cannot do everything that should be done, small specialised RFOs can target critical areas that are neglected by larger RFOs.

When it was founded, WWTF decided not to go the avenue of funding applied research or R&D as there were already a number of national and regional funders. Instead, basic research of high scientific quality that can contribute to the dynamics of the Viennese research landscape was made the primary focus. Current funding priorities target the life sciences (since 2003), interdisciplinary and applied mathematics (since 2004); a programme for information and communication technology that

focuses on basic questions of that field (since 2008), and cognitive sciences (since 2009), a newly interdisciplinary field with potential to connect areas of research in Vienna that have not yet been linked. The WWTF has a portfolio of instruments for research funding in these thematic areas, which include larger project funding as well as person-oriented funding of young group leader positions for 5-8 years ("Vienna Research Groups", abbreviated VRG), and fixed-term professorships ("Science Chairs"). Table 1 lists the programmes whose effects were investigated by the qualitative study. The majority of funds (75%) goes into Project Grants.

	Project Grants	Endowed Science Chairs	Vienna Research Groups (VRG)
Funding duration	2 to 4 years	5 years	5 years
Amount	200,000- ca. 800,000	1,5 million plus matching university funds	1,5 million plus matching university funds
Target group	all	established researchers from abroad	early career researchers from abroad

Table 1: Characteristics of investigated funding programs

2. HOW TO CHANGE A LOCAL RESEARCH LANDSCAPE

Since it is embedded in a supranational, national and regional innovation system, the support of scientific research by such small funding programmes faces a number of challenges. First, one can ask how a funding organization with limited resources can contribute to the development of a research landscape that is also addressed by much larger players. Second, as it is by no means easy to determine the success of such efforts: How can efforts to 'develop a research landscape' be evaluated?

The answer to the second question is based on the insight that any organisation should be evaluated against its mission. The qualitative study on which this section is based therefore looked for mechanisms triggered by WWTF funding that changed the Viennese basic research landscape. We found three such mechanisms.

1. CREATING LINKS AND EXPANDING STRENGTHS

A first mechanism utilises existing strength of the Viennese research landscape and links them to fields to which they have not been previously linked or expands them in new directions. This mechanism targets specific interdisciplinarity, either between basic research fields or between a fundamental and an applied field.

The WWTF issued several calls which, while thematically still sufficiently broad, asked for particular combinations of fields. The criterion interdisciplinarity is a means for combining existing research in a new way. For example, the thematic area "mathematics and ..." was created to encourage projects that apply pure mathematics in other disciplines (e.g. utilization of an innovative mathematical method in modelling and simulation). Vienna has a long tradition in formal mathematics. However, the WWTF observed that this tradition was not linked to the wide set of applications in need of such solutions. The call asked for projects to be conducted by interdisciplinary teams consisting of one applicant from mathematics and a partner from another discipline or vice versa. The pro-

jects should not just apply innovative mathematical methods but develop them further (WWTF 2008: 30).

An example for linking fundamental and applied research is the thematic call within the life sciences "linking research and patients' needs". Proposals for hypothesis-driven research aimed at strengthening links between basic research and clinical/disease-related research were invited to foster collaborations between basic scientists and clinicians.

The calls thus identify themes that fit the Viennese research landscape and create links between fields by demanding collaboration with partners from other fields in Vienna.

Well, there had to be opportunities for collaboration. This was a very, very important part of the WWTF proposal indeed, to show why this special person is able to interact with thousand and one people at the university. (Science Chair)

This selection process has to keep a balance. On the one hand, thematic areas need to be broad enough to attract enough applications. On the other hand, they have to be sufficiently focused to have a specific effect with limited funding.

In the light of limited funding, the sustainability of effects achieved with project funding is of particular interest. When new links are created by project funding: do they persist after funding runs out? The WWTF's application procedure for person-centred grants ensured that the candidates provided convincing research programmes, namely long-term plans for research.

Q: For how long is this planned? Is this now a lifelong project or ...?

A: Infinitely, yes. The fact is that you find something and then you have the next ten questions. And then you find again something, and then the next ten occur. [...]

Q: But the title [of the grant] ...

A: ... will stay for a while, I am pretty sure.

Q: A while would be decades?

A: I think so because it is a large concept. And it would need a really new finding that one suddenly moves somewhere else (LS, VRG)²

Q: These [...] research topics – for what time horizon are they designed?

A: [...] actually they are designed forever. [...] This is nothing you can finish in five years. Of course there are always intermediate goals. One wants to publish and this is why the projects are planned in small chunks. But the overall idea is planned to be very long-term. (LS, VRG)

Such long-term plans are more common for the life sciences. Many collaborations in these fields that were triggered by WWTF funding continued beyond the funded project. In these fields, in which lines of research consist of sequences of projects requiring similar combinations of expertise, WWTF funding from all three funding programmes helped researchers expanding their collaboration network.

Long-term research plans were also expected from VRG leaders in computer science as part of their application for WWTF funding. However, in computer science and mathematics the same methods can be more easily applied to new problem areas, and research programmes are likely to change much faster:

Also, the duration of the project is planned for five to eight years. However, when it approaches later phases the project description is of course vaguer because it is very difficult to plan research so far ahead. I can hardly promise what I will do in eight years because it can change very quickly. (IT, VRG leader)

Many projects in mathematics and informatics are 'isolated' in the sense that they have well-defined ad hoc – relations to other fields and don't generate new problems when solved. This is why WWTF-funded projects in these fields did initiate new links but these links were not sustained after the initial problem was solved.

Overall this WWTF Call was very important to our research group. However, you could say that the project lasted three years and we essentially did what we planned in the proposal. (MA, Project Grant)

Still, in some cases, the new projects built on ideas that emerged in prior WWTF-funded projects.

This [...] in a sense emerged suddenly. We didn't anticipate at the time [of the application] that this would become our main research area within information [...]. (IT, Project Grant)

Only researchers who could secure a permanent position and acquired grant funding from other sources as well as infrastructure support from their universities could maintain their new line of research. Several

other researchers intended to continue their research from WWTF Project Grants but couldn't due to the lack of funding.

Well, this clinical matter is of great interest to me, the clinical trial. That would be an important matter which can't be realised due to the lack of funding. (LS, Project Grant)

What happened with the project? It found some resonance [...] in the community. That's the one thing. The other thing is, of course there are follow-up questions that we would like to do. Two of them we would have taken up. One follow-up project that we tried to get funded is not yet funded, unfortunately. (IT, Project Grant)

The sustainability of research plans that evolved on the basis of WWTF funding crucially depends on continued success in acquiring external funding, which the WWTF cannot influence.

2. ATTRACTING THE (POTENTIAL) ELITE

National science policies in many countries try to attract outstanding scientists to their science systems because it is feared that the vitality of their science system is threatened by a brain drain. The underlying assumption is that the elite is critical to a nation's science base because they shape the direction of the knowledge production of their community (Mulkey 1976; Laudel 2005). Attracting the potential elite is important because they are more likely to stay and will become the new elite if provided with the right conditions. Furthermore, elite production is autocatalytic: The scientific elite selects and directs early career researchers in particularly promising research areas and increases the likelihood that these researchers become elite themselves (Mulkey 1976: 446-454, Zuckerman 1977: 99-100, Laudel 2005). Using special funding measures to attract the elite, as has been done in many countries (Laudel 2005: 377), seems to be rational. But how can they be attracted in a sustainable way, namely make them stay and having a long-term effect on a region?

The WWTF uses its funding programmes to fund professors (Science Chairs) and junior group leaders for this purpose. The candidates had to come from abroad, which made it more likely that new research areas could be brought to Vienna. Apart from that, quality and fit are the main selection criteria. This means that new research areas could only be established where "something was already there". In many research fields, the opportunities to collaborate locally are an important condition which makes a research site attractive.

There is in Vienna simply ... and there is always even more where you say, well we could, we should, let's do it. And the opportunities to collaborate are larger than what you have time for. Which is, of course, also a pleasant situation. (Science Chair)

And this [the data] you need to [analyse] cleverly and efficiently, and this is what we are currently doing a lot. And the good thing is, one has to say, it happens here because this campus is modern since people have access to these technologies the moment

1 Quotations are linked to the type of the grant the interviewee held (VRG leader, Science Chair, or Project Grant), and to the discipline (LS = Life Sciences; IT = Information and Communication Technologies; MA = Mathematics). If it is likely to make interviewees identifiable, even this information is omitted.

they are on the market. And this is why we are in the frontline of it. [...] It is good to have the people close by, to be able to talk to them and ask whether this makes sense, whether this is a meaningful result. ... Communication is really important. (Science Chair)

Many of these new collaborations of incoming elite scientists linked them to researchers in Vienna. On first glance, this looks suspicious because researchers generally search for the best collaborators regardless of their location. However, the discrepancy can be explained by the selection of applicants that fit the Viennese research landscape. Since the candidates had to come from abroad and to fit in the local research environment, new local collaborations are very likely to occur. Furthermore, local collaborations are advantageous for many research processes because they can be based on face-to-face communication. Thus, the simultaneous requirements of recruitment from abroad and local fit also made WWTF grantees in the Group Leader and Science Chairs programmes restructure their collaboration networks and tie them into the local research environment.

Candidates for professorships and for group leader positions also actively shaped their research by selecting research that both interested them and could be linked to the Viennese research landscape.

I think I visited here once more, but then there were some quite intense discussions about what [researchers in the university] were doing, what I was doing. And there we tried to find the cross links between what I thought was important in the projects that I wanted to do, indeed it is a continuation [...], the [topic 1]. [Topic 1] of course is something that has always interested me so that was more of a continuation. [Topic 2] was something that we were just a little bit getting into it. It was typically one of the things that I really wanted to do but couldn't do [at my old university abroad], I wasn't able to really focus on that. We were able to start with it but we couldn't get really much worth it. That was something that I really wanted to develop more. (LS, person-centred grant)

Since universities have to commit themselves to supporting the Chairs and group leaders (see 3.), which includes long-term investments in their respective fields, they will attempt to reap the greatest possible benefits from appointments, which includes a fit between the candidate and their own research profile. The following two quotes illustrate this 'strengthening through fit' mechanism:

I found here that there was this other group doing sort of similar kind of work that was very nice. [...] there are various groups who are doing related kind of research in Vienna. There is quite a team of this and we are doing monthly seminars. Tomorrow we have one with these groups where we meet and where we talk to each other and that is I think also very special. [...] (Science Chair)

I think the WWTF was very conscious that they employed somebody who would not come with a completely new thematic area but, while contributing a new aspect, does not create something completely new here. I find it far-sighted that they employ

people who simply fit, strengthen the whole thing and bring in even more critical mass. (VRG leader)

We have seen that the WWTF is successful in attracting excellent researcher. But how can it make sure that the researchers that were attracted through WWTF funding stay in the Vienna research landscape? Even if universities were included in the selection process of candidates (see 3.) and therefore willing to provide tenure, it is still a decision of the individual researcher whether he or she wants to continue his/her career in Vienna.

Nearly all WWTF Science Chairs were appointed to tenured positions and have remained in Vienna. One Science Chair left and took a position in Switzerland. The careful selection process and the considerable resources provided by universities create a sufficient incentive for most Science Chairs to stay on and thus ensure the sustainability of changes in the Viennese research landscape. All members of the first cohort of Vienna Research Group Leaders also received tenured positions. For the recent cohort this is still an open question. Furthermore, in contrast to the Science Chairs, the group leaders' future position come with very little funding beyond their own salary.

I mean, if I get tenure then I believe they should give me something additional, a position I guess. Otherwise I will apply somewhere else. It is not at all common that you get a professorship and nothing else than your own salary. (VRG leader)

This is a problem in fields like the biosciences, where personnel is essential for conducting research. If the WWTF made the right decision in selecting and funding these group leaders, they may receive more attractive offers from abroad when the funding period ends. In other fields where group leader positions were established, we can be more confident that at least some of the established research programmes will be continued in Vienna.

Another important condition for interviewees was having a career perspective leading to a full professorship in due time. This was considered important not only because it extends one's own resource base but also because of the authority of the position:

One is then in a position where you can begin to influence which further professors a university recruits and so on. Just from the point of the career this is a different state where you can start to shape things more closely.

The WWTF intends to attract the potential elite of a research area, and these researchers naturally want to do what the elite does, namely to shape the direction of their scientific community's research both locally and globally, as the quote above shows.

The effect of Project Grants on creating sustainable careers were less strong. Project grant holders on temporary positions did not have clear career prospects. Austrian universities provide only few tenure-track positions that lead to assistant professorships and finally to associate professorships.

Any planning of a career in Vienna is difficult. I have built some-

thing for me. I created a momentum and I would like to continue. But it is not possible that you can stay if you have achieved this and that. This is just not possible. And this is a big gap. [IT, Project Grant]

As a result, some of the WWTF grantees who received excellent offers from abroad but wanted to stay in Vienna did not find an opportunity to do so. Project grants could also have an indirect effect by increasing an early career researchers reputation and this way facilitating later achievement of tenure (however, not necessarily in Vienna).

3. 'NUDGING' RESEARCH ORGANISATIONS

The VRG Leader and the Science Chair programmes depend on the active participation of universities, which are involved in the application process and must make significant contribution in order to receive the funding for VRG Leaders respectively Science Chairs. The funding commitments were substantial in the case of Science Chairs and moderate in the case of VRG grants. A second, more important commitment expected from universities was the creation of tenured positions. While the positions of VRG Leaders and Science Chairs are funded for a fixed term, their host universities have to provide a tenure-track position for VRG Leaders, and have to establish a chair in the field of a Science Chair after their term. These commitments provide the WWTF grantees with an opportunity to move to permanent professorships and to continue their research in Vienna. This means that with its temporary funding, the WWTF leverages not only matching temporary funding from the university but also permanent investment in the fields in which the WWTF funding flows.

Since universities have to make considerable financial commitments to support the chairs and group leaders, and must commit to long-term investments in the fields of Science Chairs, they will attempt to reap the greatest possible benefits from the appointments. This is why they are concerned with the candidates' fit with their own research profile, either through strengthening existing research areas (which always coincided with the addition of a new area of expertise) or through filling important gaps in their profiles (see 2.).

The excellence of the funded professorships attracted other excellent researchers in the same research field. The initial WWTF funding acted as a trigger and created critical masses that made some areas internationally recognised strengths.

At this moment, this really is the best location for me in Europe, I think. Which also had to develop, of course. [...] This research focus [...] in Vienna is even stronger today than at the time when I came here. Some credit for that must be given to the WWTF. [It] was – with some additional luck – sufficient to establish an internationally visible research focus in Vienna. In [that area] it is simply the case that there is no better place in Europe. (Science Chair)

Even in the case of the Science Chair who left Vienna the WWTF had an impact on the research landscape because two Viennese universities decided to continue the Chair's research area by establishing their own chairs in this area (with some support of the WWTF).

Universities also created new tenure-track positions in the research areas of two Science Chairs. How such positions can be filled depends of course on the attractiveness of the Viennese research environment. At least in some cases, the Science Chairs contributed to this attractiveness.

For example, Professor X and Professor Y both came to Vienna after me. I wouldn't say it was essential that I am here but it certainly helped. The people, the famous people move where they find colleagues for whom they perhaps have some use. This is an important matter. (Science Chair)

Also at the same time someone else came, not from WWTF but from other funds. A few new groups came and this established in a way also a core of [...] sciences where I think that we have actually a very nice surrounding of the field here in Vienna at hand. [Science Chair]

The 'nudging' of research organisations, which has already proved highly successful in the case of Science Chairs and still awaits its test for VRG Leaders (the VRG programme started only in 2010), extends the WWTF's influence on the Viennese research landscape far beyond its actual funding by enrolling research organisations. It is, however, not without limits. The WWTF cannot influence the future resource base of its grantees. For VRG Leaders, this entails the possibility of moving from the position of a leader of a well-funded research group to the position of a university-financed associate professor, who receives very little resources in addition to the funding of the position itself. This is why many group leaders begin to apply for additional grants rather soon. Some WWTF grantees boosted research in their area by attracting additional funding through external grants. They also attracted people who joined their groups who brought their own funding with them.

And then the person who became my first postdoc contacted me spontaneously. He knew me from scientific meetings. He said, 'hey, maybe you are looking for a postdoc'. Then we agreed that he would come, would be funded [for some time] from WWTF, and that we would apply for a fellowship, which he received. Then another postdoc came, a very good scientist who got her [...] fellowship. And then somebody else whom I knew from meetings asked me 'are you looking for postdocs?' [...] It all worked very well indeed. (VRG)

However, success is not guaranteed. Such a prospect is particularly disheartening in the experimental fields. Science Chairs are in a similar position because the university cannot make any promises in terms of resources and personnel for the chair prior to the actual negotiations during the appointment procedure. Thus, WWTF grantees in both programmes can be confident about having a long-term perspective in terms of their position but face insecurity concerning their future research base.

3. CONCLUSIONS

The WWTF is a small RFO that is unusual in that it aims at contributing to the development of the Viennese research landscape and its fundamental research. Regional science policy actors exercise only little authority. The WWTF designs funding programmes according to thematic fit and impact of instruments, while the international scientific community is strongly involved in decisions on funding. Local scientific communities and their interests are balanced with international research communities both in the creation of thematic priorities and in the administration of calls.

In spite of its small budget, the WWTF did indeed have an impact on the research landscape of the City of Vienna. We identified three major mechanisms through which this impact was achieved. First, *creating links* and expanding strengths resulted in thematic programmes that were connected to existing research and brought in new research topics. This is in accordance with the findings of Boschma et al. (2014) that a topical link must exist to make a sustainable regional impact. Attracting the (*potential*) elite to fill the selected fields is the second mechanism. Researchers from abroad were offered generously funded tenure-track positions (junior group leader and professorial) if they met the criteria of excellent quality and fitting into the research landscape. Third, *'nudging' research organisations* means universities could receive substantial funding by the WWTF for group leaders and professors if they committed themselves to contribute matching funds and provided tenured positions for the period after the RFO funding. It 'nudged' universities towards supporting new topics because the authority in the application process was shared between the international community, the RFO and the respective university. This led to sustainable investments and in some cases to the creation of critical mass.

The WWTF is to a large extent focused on *bending academic careers*. Careers are bent thematically by focusing them on unusual cross-disciplinary relationships. At the same time, they are bent geographically by attracting and binding them to the Vienna science region.

How sustainable is this impact on the regional landscape? Since the WWTF has been operating for only a decade, this question is difficult to answer. Some tentative conclusions are possible. WWTF-funded professors are likely to continue their careers and research in Vienna. This particularly applies to those whose WWTF funding ended and who are on university-funded positions with substantial research funding. For the potential elite, the junior group leaders, sustainability is still an open question. Although all members of the first cohort have received a tenured position, these positions do not usually include substantial additional funding, which means that they might respond to attractive offers from abroad. The sustainability of careers of researchers funded by Project Grants largely depends on other actors in the region, particularly other RFOs. Thus, the more serious constraints to the WWTF's funding strategy appear to be outside its control. The WWTF provides generous funding that enables interesting research but does so only for a fixed term.

Inevitably, there are some risks to these approaches. One that has not manifested itself yet but was suggested in one interview is that with its search for unusual calls, the WWTF runs the risk of offering something

for which there is less demand, or demanding something that researchers can hardly deliver. This risk is inevitable for a small RFO that needs to capitalise on gaps in larger RFOs funding portfolios. The same logic that applies to research applies to research funding: RFOs must risk something in order to achieve something. It is not easy to see how the WWTF could achieve the same impact with a low-risk strategy.

We have provided a micro-level analysis that identified mechanisms triggered by an RFO in order to produce macro-level effects on a region's research. However, we were not able to identify these macro-level effects. While there is agreement in the science policy literature that local and regional research landscapes are important contexts for research, the empirical investigation of research landscapes, and particularly of their thematic structures, is still hampered by a lack of suitable methods and data. It is still not clear how quality, thematic structure, epistemic diversity and coherence of a regional research landscape can be measured. The approach of choice would be bibliometric methods, possibly combined with network analysis. However, such an analysis would require the generation of complex data sets and is thus a question of the cost-benefit ratio. But even when we will be able to measure an RFO's impact on a regional research landscape, there will be substantial impacts which will defy quantitative detection as the dynamics of research cannot be represented in rather fixed denominations and categories of research output.

REFERENCES

- Boschma, Ron, Gaston Heimeriks, and Pierre-Alexandre Balland,** 2014. Scientific knowledge dynamics and relatedness in biotech cities. *Research Policy* 43: 107-114.
- Braun, Dietmar,** 1993. Who Governs Intermediary Agencies? Principal-Agent Relations in Research Policy-Making. *Journal of Public Policy* 13: 135-162.
- Braun, Dietmar,** 1998. The role of funding agencies in the cognitive development of science. *Research Policy* 27: 807-821.
- Braun, Dietmar, and David H. Guston,** 2003. Principal-agent theory and research policy: an introduction. *Science and Public Policy* 30: 302-308.
- Caswill, C.,** 2003. Principals, agents and contracts. *Science and Public Policy* 30: 337-346.
- Costa, Mariona, Irwin Feller, Martin Grötschel, Grit Laudel, Gunnar Öquist, and Oliver Stock,** 2014. WWTF Impact Evaluation 2013/2014: Report by the International Review Panel. Wien: WWTF, http://wwtf.at/upload/WWTF_impacteval_PanelReport.pdf.
- Gertner, Drew, and Bart AG Bossink,** 2015. The evolution of science concentrations: The case of Newcastle Science City. *Science and Public Policy* 42: 121-138.
- Gläser, Jochen, and Grit Laudel,** 2015. A Bibliometric Reconstruction of Research Trails for Qualitative Investigations of Scientific Innovations. *Historical Social Research - Historische Sozialforschung* 40: 299-

330.

Gläser, Jochen, and Grit Laudel, 2016. How does governance change research content? Two approaches waiting for integration. *European Journal of Sociology* 57: forthcoming.

Guston, David H., 1996. Principal-Agent Theory and the Structure of Science Policy. *Science and Public Policy* 23: 229-240.

Hoekman, Jarno, Koen Frenken, and Robert J.W. Tijssen, 2010. Research collaboration at a distance: Changing spatial patterns of scientific collaboration within Europe. *Research Policy* 39: 662-673.

Inhaber, H., 1974. Scientific cities. *Research Policy* 3: 182-200.

Laudel, Grit, 2005. Migration Currents among the scientific elite. *Minerva* 43: 377-395.

Laudel, Grit, 2013. An inDepth case study of selected WWTF impacts. Technical University Berlin, Institute of Sociology, http://wwtf.at/upload/WWTF_impacteval_CasestudyLaudel.pdf.

Morris, Norma, 2003. Academic Researchers as 'agents' of science policy. *Science and Public Policy* 30: 359-370.

Mulkay, Michael, 1976. The Mediating Role of the Scientific Elite. *Social Studies of Science* 6: 445-470.

Rip, Arie, 1994. The Republic of Science in the 1990s. *Higher Education* 28: 3-32.

Rip, Arie, 2002. Regional Innovation Systems and the Advent of Strategic Science. *The Journal of Technology Transfer* 27: 123-131.

Shove, Elizabeth, 2003. Principals, agents and research programmes. *Science and Public Policy* 30: 371-381.

Tripl, Michaela, and Gunther Maier, 2010. Knowledge spillover agents and regional development. *Papers in Regional Science* 89: 229-233.

Van den Daele, Wolfgang, Wolfgang Krohn, and Peter Weingart, 1977. The Political Direction of Scientific Development. E. Mendelsohn, P. Weingart and R.D. Whitley (eds.), *The Social Production of Scientific Knowledge*. Dordrecht: Reidel, 219-242.

van der Meulen, Barend, 1998. Science policies as principal-agent games: Institutionalization and path dependency in the relation between government and science. *Research Policy* 27: 397-414.

van der Meulen, Barend, 2003. New roles and strategies of a research council: intermediation of the principal-agent relationship. *Science and Public Policy* 30: 323-336.

Van der Meulen, Barend, and Arie Rip, 1998. Mediation in the Dutch science system. *Research Policy* 27: 757-769.

Van Geenhuizen, Marina, and Peter Nijkamp, 2012. *Creative knowledge cities: myths, visions and realities*. Edward Elgar Publishing.

WWTF [Wiener Wissenschafts-, Forschungs- und Technologiefonds], 2008. Background Paper WWTF Review Panel. http://www.fteval.at/upload/BackgroundPaper_WWTF.pdf.

WWTF Office, 2014. Impact Evaluation 2013/14: Selfevaluation Report to International Review Panel. Wien: WWTF, http://wwtf.at/upload/WWTF_impacteval_SelfevalReport.pdf.

Zuckerman, Harriet A., 1977. *Scientific Elite: Nobel Laureates in the United States*. New York: Free Press.

AUTHORS

GRIT LAUDEL

Technische Universität Berlin

E: grit.laudel@tu-berlin.de

MICHAEL STRASSNIG

Wiener Wissenschafts-, Forschungs- und Technologiefonds WWTF

E: michael.strassnig@wwtf.at

ZWISCHENEVALUIERUNG DER VOM BMWF BEAUFTRAGTEN REGIONALEN KONTAKTSTELLEN – ODER DIE EVALUIERUNG VON INTERMEDIÄREN STRUKTUREN IN IHREM KONTEXT

BARBARA GOOD

AUSGANGSLAGE

Die insgesamt vier vom damaligen Bundesministerium für Wissenschaft und Forschung (BMWF) finanzierten Regionalen Kontaktstellen (RKS) waren Teil des Beratungs- und Betreuungsnetzwerkes, das das BMWF seit dem Beitritt Österreichs zu den EU-Forschungsrahmenprogrammen eingerichtet hatte. Das Ziel der RKS war die Betreuung der Forscher/innen in der jeweiligen Region zur Optimierung der österreichischen Beteiligung an den EU-Forschungsprogrammen und die Sicherstellung einer bestmöglichen Information, Beratung und Betreuung von Wissenschaft, Wirtschaft und Verwaltung in Bezug auf das EU-Forschungsrahmenprogramm und andere europäische Programme in diesem Bereich.

Im Herbst 2012 beauftragte das BMWF Technopolis mit der Evaluierung der RKS. Die Evaluierung hatte zum Ziel, die Dienstleistungen sowie die Struktur der Regionalen Kontaktstellen im Hinblick auf ihre Effektivität und Effizienz zu analysieren und Empfehlungen für eine zukünftige Optimierung sowie - unter Berücksichtigung der auf EU-Ebene zu erwartenden Veränderungen - Grundlagen für allfällige Folgebeauftragungen zu erarbeiten. Bewertet wurde der Zeitraum der RKS-Aktivitäten zwischen 2009 und 2012. Die Evaluierung bezog sich also noch auf das alte EU-Rahmenprogramm (FP7).

Alle vier RKS waren in die jeweilige regionale Innovationsagentur eingebettet, und in keiner der regionalen Innovationsagenturen bildete die Regionale Kontaktstelle eine eigene Abteilung. Bei den regionalen Innovationsagenturen handelte es sich zum Zeitpunkt der Evaluierung um die Standortagentur Tirol (SAT), die damalige CATT Innovation Management GmbH (Oberösterreich), die Steirische Wirtschaftsförderungsgesellschaft (SFG) sowie die ITG Innovations- und Technologietransfer Salzburg GmbH. Auch von den Kund/innen wurden die RKS nicht als eigenständige Organisationseinheit wahrgenommen, sondern als die EU-Förderberatung in der jeweiligen Agentur. Insofern haben wir mit den RKS Einheiten evaluiert, die als solche in den Agenturen nicht sichtbar waren.

Die vier RKS verfügten über eine unterschiedliche Entstehungsgeschichte, waren auf unterschiedliche Art und Weise in die jeweilige regionale Innovationsagentur eingebettet, verfügten über teilweise unterschiedliche Aufgaben (insbesondere CATT) und waren in einem unterschiedliches wirtschaftlichen Umfeld tätig. Dies bedingte einen methodischen Zugang, der diesen Unterschieden gerecht wurde. Für die Evaluierung kamen dementsprechend Dokumentenanalysen (namentlich Tätigkeitsberichte, Kostenaufstellungen) und Sekundärdatenauswertungen (Kundendaten der RKS), Interviews im Rahmen eines Besuchsprogramms in den Bundesländern, Stakeholderinterviews sowie eine standardisierte Onlineumfrage bei Kund/innen der RKS zum Einsatz. Effektiv bedeutete dies, dass wir im Rahmen der Zwischenevaluierung vier Mini-Evaluierungen vornahmen und dass die Vergleichsmöglichkeiten zwischen den RKS limitiert waren, wenn wir nicht Äpfel mit Birnen vergleichen wollten.

DIE VIER REGIONALEN KONTAKTSTELLEN IM EINZELNEN

DIE REGIONALE KONTAKTSTELLE IN DER STEIRISCHEN WIRTSCHAFTSFÖRDERUNG (SFG)

In der Steiermark wurde der Verein „APS – European Programmes for Technologies and Training“, welche bis dahin die RKS-Funktion in der Steiermark und in Kärnten ausübte, mit 1. Jänner 2009 in die Steirische Wirtschaftsförderung (SFG) integriert. Diese Integration repräsentierte einen kompletten Neuanfang: Die SFG bekam keine Unterlagen der APS und übernahm auch keine Mitarbeiter/innen. Damit war 2009 der Aufbau von Know-how eine der vorrangigen Aufgaben der steirischen RKS.

Die Integration in die SFG ermöglichte es der RKS, einen ganzheitlichen Beratungsansatz zu verfolgen. Dabei ging es nicht darum, den Kunden um jeden Preis von einer Teilnahme am FP7 zu überzeugen, sondern

seine Bedürfnisse in den Mittelpunkt zu stellen und ihn bei Bedarf an andere Förderungen weiterzuleiten. In diesem Zusammenhang wurde oft von „Wegweiserfunktion“ gesprochen, bei der es darum geht, den Kunden zur richtigen Förderung hinzuführen – sei dies zum Rahmenprogramm oder anderen europäischen Programmen, sei dies zu nationalen oder regionalen Förderungen.

In der Beratung bestand eine klare Arbeitsteilung der RKS in der SFG mit der FFG/EIP. Die RKS in der SFG beriet Antragsteller/innen nur in der

Antragsphase. Proposal Checks¹ unternahm sie keine. Auch die Projektbegleitung ließ sie weg. Beides – Proposal Checks und Projektbegleitung – überließ sie der FFG/EIP. Die SFG übernahm auch relativ wenige Detailberatungen; auch diese überließ sie der FFG. Dies erschien sinnvoll, da die National Contact Points (NCPs) in der FFG über mehr Detailwissen verfügten als die RKS, die ein eher breit gefächertes Wissen (inkl. über andere regionale und nationale Förderungen) besaßen.

Calls 2010/2011	Partnerprojekte (n=522)	Koordinatorprojekte (n=141)
Beratung durch	%	%
FFG und/oder SFG	48%	73%
- davon nur FFG	39%	72%
- davon nur SFG	26%	30%
keine Beratung	52%	27%

Tabelle 1: Beratung von RP-Antragsteller/innen durch FFG und/oder SFG
Quelle: INNO-RKS. Alle Antragsteller/innen (auch von Universitäten).

Tabelle 1 zeigt, wieviele Antragsteller/innen in der Steiermark und Kärnten von der FFG und/oder der SFG beraten wurden. Die Zahlen beziehen sich auf die Calls aus den Jahren 2010/2011. Die SFG betreute gut ein Viertel (26%) der Antragsteller/innen, die als Partner in einem Konsortium teilnehmen, verglichen mit knapp einem Drittel (30%) der Antragsteller/innen, die als Koordinator einreichen. Verglichen mit der SFG beriet die FFG indes sowohl mehr Partner als auch mehr Koordinatoren.

Die Anzahl Beratungen der RKS nahm ab der Integration in die SFG (2009) kontinuierlich zu. Während 2009 noch rund 25% der Beratungen an Universitäten stattfanden, sank dieser Anteil 2010 auf 10%.² Dies widerspiegelte die Implementation der Arbeitsteilung mit den Forschungsservices. Während die Forschungsservices ihre eigenen Angehörigen berieten, betreute die SFG Unternehmen, außeruniversitäre Forschungseinrichtungen, Intermediäre (z.B. Cluster) und Fachhochschulen. Bei einschlägigen Veranstaltungen arbeitete die SFG indes intensiv mit den Forschungsservices der Universitäten zusammen.

DIE REGIONALE KONTAKTSTELLE IN DER CATT

Wie bei der SFG bildete die Regionale Kontaktstelle in der damaligen CATT keine separate Abteilung, sondern war Teil des Bereichs Förder-Management. Das Förder-Management beschäftigte zum Zeitpunkt der Evaluierung rund 10 Mitarbeiter/innen, wovon eine Mehrheit für die RKS tätig war³. Wie die SFG verfolgte auch die CATT einen ganzheitlichen Beratungsansatz und leitete bei Bedarf einen Kunden oder eine Kundin an eine ihm/ihr angepasste Förderung weiter („Wegweiserfunktion“).

Die RKS in der CATT verfolgte in ihren Tätigkeiten jedoch ein anderes Modell als die anderen RKS. Eine klare Arbeitsteilung mit der FFG/EIP war nicht sichtbar. Vielmehr war das Modell der CATT das einer „Mini-FFG“. Die CATT beriet über alle Projektphasen des Europäischen Rahmenprogramms und anderer europäischer Programme hinweg. Dies betraf die Beratung und Betreuung zu Fragen der Projektentwicklung, die Einreichung sowie die Abwicklung und das Reporting. Die CATT bot auch Proposal Checks an.

Calls 2010/2011	Partnerprojekte (n=196)	Koordinatorprojekte (n=59)
Beratung durch	%	%
FFG und/oder CATT	53%	75%
- davon nur FFG	43%	69%
- davon nur CATT	26%	44%
keine Beratung	47%	25%

Tabelle 2: Beratung von RP-Antragsteller/innen durch FFG und/oder CATT
Quelle: INNO-RKS. Alle Antragsteller/innen (auch Universitäten).

¹ Die Überprüfung eines Antrags zum EU-Rahmenprogramm im Rahmen eines detaillierten „Proposal Checks“ ist eine Standarddienstleistung des Bereichs EIP der FFG und wurde auch von einem Teil der RKS angeboten.

² Tätigkeitsberichte 2009 und 2010 der RKS/SFG.

³ Kostenplan 2012

Auffallend war der relativ hohe Anteil an Koordinator/innen, welche die CATT betreute (Tabelle 2). Dies stand im Einklang mit CATTs Aufgabenbereich, der die Betreuung von Universitätsangehörigen, die oft als Koordinatoren einreichen, miteinschloss. Im Gegensatz zur SFG und den anderen RKS beriet die CATT auch die Angehörigen der Universität Linz, was auch der laufenden Beauftragung entsprach. Die Angehörigen der Universität Linz machten gar die Hauptkundengruppe aus. Genauso wie bei der SFG betreute indessen die FFG/EIP in Oberösterreich sowohl mehr Partner als auch mehr Koordinatoren als die CATT (Tabelle 2).

DIE REGIONALE KONTAKTSTELLE IN DER ITG

Mit Beginn 2011 wurde der Innovationsservice der Wirtschaftskammer Salzburg in die ITG Salzburg integriert. Damit wurde im Land Salzburg die Innovationsförderberatung unter einem Dach – demjenigen der ITG – gebündelt. Die Zusammenlegung der Beratungsstellen vollzog sich nicht nur räumlich, sondern insbesondere durch die Einführung eines neuen, ganzheitlichen Beratungsansatzes.

Die ITG verfolgte eine ähnliche Arbeitsteilung mit der FFG/EIP wie die SFG. Die Erstberatungen und allgemeine Beratungen zur Heranführung an europäische Programme fanden beauftragungskonform bei der ITG statt, bei Detailfragen verwies man die Kund/innen an die FFG/EIP bzw. klärte diese Fragen selber bei FFG/EIP ab.

Auch mit dem Forschungsservice der Universität Salzburg (Abteilung für Forschungsförderung) bestand eine klare Arbeitsteilung. Der Forschungsservice betreute die Angehörigen der Universität und die ITG betreute die außeruniversitären Kunden, d.h. Angehörige der Fachhoch-

schule, von außeruniversitären Forschungseinrichtungen und Firmen. Daneben arbeitete man seit langem bei der Konzeption und Organisation von Veranstaltungen, Workshops, Schulungen und Tagungen zusammen.

Salzburg verfügt über eine strukturelle Schwäche in Bezug auf forschungsintensive Betriebe, was den Pool an potentiellen RP-Teilnehmern unter den Unternehmen beschränkt. Dies widerspiegelte sich in den Beratungszahlen vor Einführung des ganzheitlichen Beratungsansatzes, als nur wenige Beratungen ausschließlich zum Europäischen Rahmenprogramm oder anderen europäischen Programmen durchgeführt wurden. Nach der Einführung des ganzheitlichen Beratungsansatzes nahm die Anzahl an Beratungen indes sehr stark zu. 2012 wurden laut Bericht der ITG an den Auftraggeber BMWF zehn Mal mehr Beratungen durchgeführt als 2009 (+1000%). Diese Zunahme hing damit zusammen, dass die RKS-Berater/innen nicht nur zu europäischen Programmen, sondern auch zu regionalen und nationalen Programmen berieten. Dies bedeutete aber auch, dass die Beratungen de facto inhaltlich oft keinen Zusammenhang mit den Europäischen Rahmenprogramm und anderen europäischen Programmen aufwiesen, obwohl sie in der Gesamtstatistik (auch) als EU-relevante Beratungen gezählt wurden.

Die Wichtigkeit der Wegweiserfunktion bzw. der allgemeinen Beratung zu (europäischen) Programmen lässt sich auch aus Tabelle 3 herauslesen. Die Tabelle zeigt aber auch, dass die ITG nur sehr schwach in die konkrete Beratung von RP-Antragsteller/innen eingebunden war und Antragsteller/innen aus Salzburg nur in einer kleinen Minderheit der Fälle die Dienste der ITG nutzten.

Calls 2010/2011	Partnerprojekte (n=102)	KoordinatorInnenprojekte (n=34)
Beratung durch	%	%
FFG und/oder ITG	56%	76%
- davon nur FFG	54%	76%
- davon nur ITG	2%	0%
keine Beratung	44%	34%

Tabelle 3: Beratung von RP-Antragsteller/innen durch FFG und/oder ITG
Quelle: INNO-RKS. Alle Antragsteller/innen (auch von Universitäten).

DIE REGIONALE KONTAKTSTELLE IN DER STANDORTAGENTUR TIROL (SAT)

Mit Jänner 2011 wurde der Verein „Büro für Europäische Programme“ (BEP) mitsamt seinen Mitarbeiter/innen in die Standortagentur Tirol (SAT) integriert. Dadurch sollte Kontinuität bei der Betreuung der Tiroler Universitäten und Unternehmen im 7. Rahmenprogramm gewährleistet werden. Bis Mitte 2011 verließen jedoch zwei der vom BEP übernommenen Mitarbeiter/innen die SAT. Dadurch entstand eine Lücke in der Betreuung des Rahmenprogramms bei der SAT. Erst durch den Neuzugang einer Mitarbeiterin Ende Oktober 2011 konnte die Betreuung des Rahmenprogramms in Tirol wieder voll aufgenommen werden. Wie bei den anderen RKS ermöglichte die Integration der EU-Förderberatung eine ganzheitliche Beratung.

Es war eine klare Arbeitsteilung zwischen SAT und FFG/EIP sichtbar. Die SAT beriet in erster Linie im Vorfeld der Antragsphase bzw. in der Antragsphase. Dann gab sie den Antragsteller/die Antragstellerin an die FFG weiter, teilweise auch aus Ressourcengründen. Die RKS in der SAT ging auch proaktiv auf Unternehmen zu und informierte sie zum EU-Rahmenprogramm und anderen europäischen Programmen.

Es herrschte eine strikte Arbeitsteilung zwischen der RKS in der SAT und den Forschungsservices der Universität Innsbruck bzw. der Meduni Innsbruck. Schon vor der Integration des BEP in die SAT im Jänner 2011 betreute der Forschungsservice der Universität Innsbruck seine Angehörigen in Bezug auf europäische Programme selber. Der Forschungsservice der Meduni Innsbruck verstärkte seine Kapazitäten in Bereich

der europäischen Programme mit der Einstellung eines ehemaligen BEP-Mitarbeiters Mitte 2011. Seither betreut die Meduni ihre Angehörigen ebenfalls selber. Es gab auch kaum gemeinsame Veranstaltungen zwischen der SAT und den Forschungsservices, wie man dies beispielsweise aus Graz und Salzburg kannte.

Calls 2010/2011	Partnerprojekte (n=251)	Koordinatorprojekte (n=77)
Beratung durch	%	%
FFG und/oder SAT (BEP)	46%	71%
- davon nur FFG	33%	65%
- davon nur SAT (BEP)	20%	21%
keine Beratung	54%	29%

Tabelle 4: Beratung von RP-Antragssteller/innen durch FFG und/oder SAT (BEP)
Quelle: INNO-RKS. Alle Antragssteller/innen (auch von Universitäten).

Wie bei den anderen RKS betreute die FFG/EIP im Tirol sowohl mehr Partner als auch mehr Koordinatoren als die SAT (Tabelle 4). Bei den Koordinatoren ist dieser Anteil bedeutend höher (65% verglichen mit 21%).

Schließlich ließ sich bei der SAT eine hohe Konzentration an Beratungen bei wenigen Organisationen feststellen, was nach einer Verbreiterung der Kundenbasis rief. Eine bessere Sichtbarkeit der SAT in Bezug auf europäische Programme erschien als ein erster Schritt, wurde die SAT in ihrer Außenwirkung doch kaum mit den europäischen Programmen in Verbindung gebracht.

RKS-ÜBERGREIFENDE ERGEBNISSE

SCHWINDEN DER ZIELGRUPPEN DER RKS-BERATUNGEN

Die Evaluierung zeigte, dass bei den Beratungen eine klare Arbeitsteilung zwischen den RKS und den Forschungsservices der Universitäten bestand, wonach die Forschungsservices die Angehörigen ihrer Universität betreuten und die RKS die restlichen Interessent/innen am EU-Rahmenprogramm und anderen europäischen Programmen. Bei der Konzeption und Organisation arbeiteten die RKS und die Forschungsservices der Universitäten hingegen oft zusammen, insbesondere an den Standorten Graz und Salzburg.

Eine Ausnahme von der Arbeitsteilung zwischen RKS und Forschungsservices bildete die CATT, die auch die Angehörigen der Universität Linz betreute und dort auch Workshops und Trainings anbot. Die Resultate dieser und früherer Evaluierungen⁴ legten jedoch nahe, dass es sinnvoller wäre, wenn die Universitätsangehörigen vom eigenen Forschungsservice beraten würden, denn dieser kennt die spezifischen Gegebenheiten der Universität.

Hauptkund/innen der RKS waren in der Folge Unternehmen, insbesondere KMU. KMU sind aber keine einfache Kundengruppe. Denn erstens ist die Gruppe an KMU, die forschungsbasiert genug sind, um an europäischen Programmen teilzunehmen, begrenzt. Und zweitens nehmen KMU meist als Partner in EU-Projekten teil und können neben FFG/EIP auch auf eine Betreuung durch den Koordinator des Projektes zurückgreifen. Auch Angehörige der außeruniversitären Forschungseinrichtungen, Fachhochschulen, Privatuniversitäten und von Intermediären gehörten zu den Kund/innen der RKS. Indes zeigten die Interviews mit ausgewählten Kund/innen, dass gerade Wissenschaftler/innen die RKS weniger häufig brauchen, weil sie die europäischen Programme meist gut kennen und sich bei Fragen direkt an die FFG/EIP wenden.

Der Wegfall der Universitäts-Angehörigen als Hauptkundengruppe der RKS widerspiegelte sich kaum in den Budgets der RKS, die von der vorherigen Beauftragung zur laufenden mehr oder weniger konstant geblieben waren. Anstelle der Beratung von Universitätsangehörigen beauftragte das BMWF die RKS mit strategischen Aufgaben (z.B. proaktive Beiträge zur Ausarbeitung regionaler FTI-Strategien oder an Konsultationsprozessen zu H2020). Die CATT betreute mit demselben Budget wie die SAT und die SFG⁵ auch die Angehörigen der Universität Linz, übernahm aber wie alle anderen RKS ebenfalls strategische Aufgaben.

Insgesamt machten die strategischen Aufgaben keinen sehr großen Anteil der Arbeit der RKS aus. Wie stark sich die RKS in strategische Prozesse einbringen konnten, hing auch davon ab, inwiefern sie und ihre Agentur in die strategischen Prozesse eingebunden waren und ob überhaupt strategische Prozesse stattfanden.

WEGWEISERFUNKTION ALS KONSTITUIERENDES MERKMAL DER RKS

Die Integration der RKS in die Landesagenturen hatte ohne Zweifel eine Professionalisierung mit sich gebracht. Dabei passten die RKS gut

⁴ Erik Arnold et al., Evaluation of Austrian Support Structures for FP 7 & Eureka and Impact Analysis of EU Research Initiatives on the Austrian Research & Innovation System, Technopolis, 2010.

⁵ Die RKS in der ITG verfügte über ein geringeres Budget als die anderen RKS, weil die ITG vor Einführung des ganzheitlichen Beratungsansatzes Anfang 2011 signifikant weniger Beratungen als die anderen RKS durchführte.

ins Portfolio der Landesagenturen. Die Integration in die Landesagenturen verbesserte nicht nur die Verankerung der regionalen Kontaktstellen im regionalen Innovationssystem, sie ermöglichte auch strategische Beratungen zum Rahmenprogramm und anderen europäischen Programmen mit Wegweiserfunktion. Dabei zeigte sich im Gesamtbild, dass die RKS auftragsgemäß ihre Funktion als Wegweiser mit niederschwelligem Beratungsangebot weitgehend erfüllten.

Indes gab es auch Probleme mit der Wegweiserfunktion. Bei Beratungen war teilweise schwierig zu eruieren, wie stark die europäischen Programme thematisiert werden. Bei der SAT hatten rund 50% der Beratungen eine Wegweiserfunktion, wonach beratene Firmen an nicht-europäische Programme weitergeleitet wurden. Bei ITG waren es angesichts des rasanten Anstiegs der Anzahl Beratungen nach der Einführung des ganzheitlichen Beratungsansatzes noch mehr. Ein stattdessen Anteil von Kunden – je nach RKS, konservativ geschätzt, zwischen rund 7% und knapp 28% – gab in der Umfrage an, keine Informationen zu EU-Förderprogrammen erhalten zu haben. Es stellte sich vor diesem Hintergrund einerseits die Frage nach der Relevanz der RKS-Beratungen für die Hauptkundengruppe, nämlich die Unternehmen, aber auch nach der Qualität der Kundendatenbanken bei den RKS.

Insgesamt war die Zufriedenheit der Kund/innen mit der erbrachten Beratungsleistung bei allen RKS aber sehr hoch. Besonders geschätzt wurde die regionale Nähe und die Coaching- und Wegweiserfunktion, d.h. die kurzfristige Möglichkeit für alle Förderfragen jemand anzusprechen, der evtl. auch dem Unternehmen eine Visite abstattet. Auch in Zeiten der digitalen Kommunikationsmittel stellte sich persönlicher Kontakt als wichtig heraus, vor allem für unerfahrene Antragsteller/innen.

Das mögliche Wirkungsspektrum der RKS-Aktivitäten war jedoch beschränkt, einerseits auf Grund der Kundenstruktur – Stichwort: Konzentration auf KMU – und andererseits auf Grund der Tätigkeiten von FFG/EIP für erfahrenere Teilnehmer/innen am EU-Rahmenprogramm. In der Summe ergab sich ein Bild, dass ein etwaiger Mehrwert einer RKS gegenüber FFG/EIP vor allem in einer allgemeinen Förderberatung und einer Coaching- und Wegweiserfunktion in diesem Kontext bestand, die sich vor allem an Einsteiger/innen in das Förderwesen richtete. Hingegen bot die reine EU-Förderberatung nicht wirklich einen Mehrwert. Des Weiteren hatten die RKS bei Awareness-Maßnahmen zu europäischen Programmen (z.B. beim proaktiven Ansprechen von Unternehmen) eine wichtige Rolle zu spielen, da hier ein Wissen um die regionalen Player unerlässlich ist. Diese Rolle konnte die FFG/EIP weniger gut übernehmen, da ihr die Akteure in den regionalen Innovationssystemen in geringerem Umfang bekannt waren.

SCHLUSSFOLGERUNGEN

Die Resultate der Evaluierung legten nahe, dass angesichts der relativ niedrigen Relevanz der RKS-Leistungen für viele Kund/innen und einer überschaubaren Zielgruppe ein Wegfall der RKS-Beratungsleistungen keinen Flurschaden anrichten würde. Die Interviews zeigten, dass

die Interessent/innen an EU-Programmen bei allfälligen Fragen und Problemen ihren Weg zur FFG/EIP finden würden. Der Vorteil der RKS lag in ihrer regionalen Nähe und ihrer regionalen Verankerung, ihrer Fähigkeit zu einer Zusammenschau aller Förderprogramme, was letztlich ihr Mehrwert war – und nicht die EU-Förderberatung an sich.

Das Ergebnis, dass die RKS einerseits ihre Aufgabe erfüllten, andererseits aber mit einer geschrumpften Zielgruppe konfrontiert waren, für die EU-Förderprogramme nur zum Teil von hoher Relevanz sind, machte es, mit wenigen Ausnahmen, schwierig, eindeutige Maßnahmenempfehlungen auszusprechen. Die wesentliche Frage, die sich stellte, war, ob der erbrachte und mögliche Nutzen den Mitteleinsatz rechtfertigte.

Zusätzlich wurde die Situation dadurch kompliziert, dass die Maßnahmenempfehlungen zweckmäßigerweise auch aktuelle und absehbare zukünftige Entwicklungen berücksichtigen mussten. Zum Zeitpunkt der Evaluierung bestanden jedoch erhebliche Unsicherheiten insbesondere hinsichtlich der Ausgestaltung des zukünftigen EU-Forschungsrahmenprogramms Horizon 2020 oder der zukünftigen Ausrichtung der FFG/EIP. Vor diesem Hintergrund arbeitete das Evaluierungsteam drei Alternativszenarien und ein Zusatzszenario aus:

- Alternativszenario 1: RKS-Beauftragung des BMWF auslaufen lassen
- Alternativszenario 2: Minimalkompetenz zu europäischen Programmen in Landesagenturen
- Alternativszenario 3: Beibehaltung des Status Quo mit Modifikationen und Anpassung an zukünftige zusätzliche Herausforderungen
- Zusatzszenario 4: strategischer Dialog mit Bundesländern

Das BMWFW entschied sich für das Szenario 1 und ließ die Beauftragung Mitte 2014 auslaufen. Dies war nicht zuletzt eine Folge der vorliegenden Evaluierung sowie eines Prüfberichts des Rechnungshofs über die nationalen Maßnahmen zum 7. EU-Rahmenprogramm, der die Einstellung der Finanzierung der Regionalen Kontaktstellen durch das BMWF eindeutig empfahl.⁶

AUTORIN

BARBARA GOOD

Technopolis Group Austria

E: barbara.good@technopolis-group.com

EVALUIERUNG VON „YOUNG SCIENCE“ – EINEM PROJEKT AN DER SCHNITTSTELLE ZWISCHEN SCHULE UND WISSENSCHAFT

CAROLINE MANAHL UND SUSANNE DOBNER

Das Zentrum für soziale Innovation (ZSI) wurde Ende 2014 mit der Evaluierung von „Young Science“ beauftragt. Dabei handelt es sich um ein vom Bundesministerium für Forschung, Wissenschaft und Wirtschaft (BMWFW) finanziertes und von der österreichischen Agentur für internationale Mobilität und Kooperation in Bildung, Wissenschaft und Forschung (OeAD) umgesetztes Projekt an der Schnittstelle zwischen Schule und Wissenschaft. Das Projekt findet sich so nicht bloß an einer Schnittstelle zwischen zwei Bildungssystemen; die Tätigkeiten von Young Science grenzen auch an die Aufgabenbereiche zweier Bundesministerien.

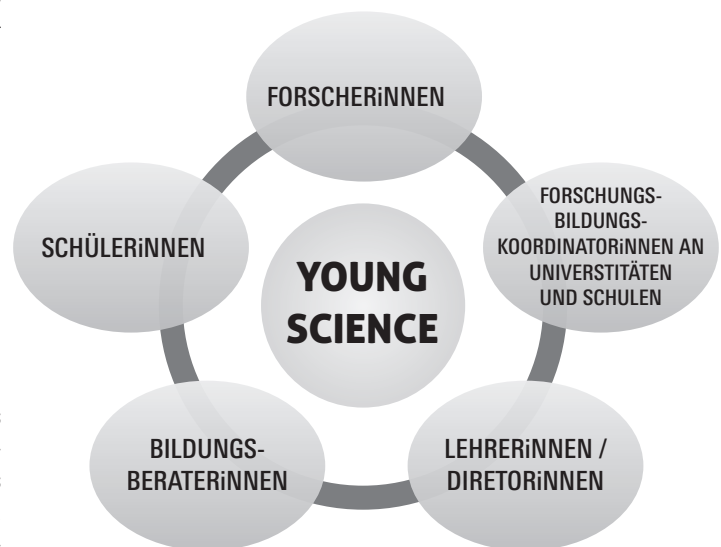
ZIELE UND AUFGABEN VON „YOUNG SCIENCE“

Seit 2011 wird das Projekt Young Science mit dem Ziel der Förderung der Zusammenarbeit von Schulen und Wissenschaft umgesetzt. Ziel des Projekts ist es, durch die Einrichtung einer Netzwerk- und Serviceplattform die Synergieeffekte bestehender und zukünftiger Initiativen des BMWFW an der Schnittstelle zwischen Schule und Forschung besser zu nutzen, die Kooperation zwischen sekundären und tertiären Bildungseinrichtungen zu fördern und Anknüpfungspunkte zum Bundesministerium für Bildung und Frauen sowie zu anderen Ressorts zu ermöglichen.

Primäre Zielgruppen von Young Science sind LehrerInnen, BildungsberaterInnen, SchülerInnen und ForscherInnen. Aber auch andere Personen, die an der Schnittstelle zwischen Schule und Wissenschaft arbeiten, werden von Young Science angesprochen (z.B. Forschungs-Bildungs-KoordinatorInnen an österreichischen Universitäten). In den vergangenen Jahren hat sich Young Science von einer Informations- und Servicestelle zunehmend zu einer Plattform weiter entwickelt, die eigene Projekte und Initiativen umsetzt.

Zentrale Aktivität von Young Science ist die Vernetzung verschiedener StakeholderInnen an der Schnittstelle von Schule und Wissenschaft. Dies geschieht durch Veranstaltungen, aktive Kontaktabbau und Öffentlichkeitsarbeit (Young Science betreibt eine Webseite¹ und gibt einen Newsletter heraus). Zudem verleiht Young Science mit dem „Young Science Gütesiegel“ eine Auszeichnung an Schulen mit einer he-

rausragenden Kooperation mit wissenschaftlichen Einrichtungen, setzte Schulprojekte und Praktikainitiativen zum Thema Nachhaltigkeit um und betreibt neuerdings eine Plattform, die SchülerInnen bei der Themenfindung für ihre vorwissenschaftlichen Arbeiten unterstützt.



Grafik1: Zielgruppen von Young Science

METHODISCHE HERANGEHENSWEISE

Bei der Evaluierung von „Young Science“ wurden zwei Schwerpunkte gesetzt – einer auf Vernetzungsaktivitäten und einer auf Projekte, mit deren Umsetzung Young Science vom BMWFW beauftragt wurde. Dabei kamen qualitative und quantitative Methoden der Datensammlung und -auswertung zur Anwendung.

Neben den Auswertungen von Projektanträgen, Berichten, Webstatistiken und Veranstaltungsdaten wurden im Dezember 2014 zwölf VertreterInnen der Young Science-Zielgruppen und MitarbeiterInnen von

Young Science im Rahmen von Workshops und Interviews befragt. Ziel war es, Einsichten in den Nutzen von „Young Science“ und die Motivation zur Teilnahme an Young Science-Aktivitäten zu erlangen, Barrieren für eine bessere Zusammenarbeit von Schulen und Forschungseinrichtungen zu identifizieren und mögliche Entwicklungspotenziale (z.B. im Bereich Citizen Science) auszuloten. In diesem Artikel wird insbesondere auf jene Evaluierungsergebnisse eingegangen, die in engem Zusammenhang mit der Positionierung des Zentrums an der Schnittstelle zwischen Schule und Wissenschaft stehen.²

INSTITUTIONELLE UND ORGANISATORISCHE RAHMENBEDINGUNGEN FORDERN VON YOUNG SCIENCE EINE HOHE FLEXIBILITÄT UND SERVICEORIENTIERUNG

Eine zentrale Aufgabe von Young Science ist die Vernetzung. Durch Young Science sollen Kontakte hergestellt und Kooperationen zwischen Personen und Institutionen gefördert werden, die ansonsten nicht bestehen. Eine Voraussetzung für die bessere Vernetzung von Schule und Wissenschaft sind an erster Stelle interessierte Personen an Schulen und Forschungseinrichtungen.

Die Evaluierungsergebnisse machen Gemeinsamkeiten und Unterschiede, die den Erfolg von Vernetzungsaktivitäten beeinflussen, für verschiedene Zielgruppen sichtbar. Durch die Workshops und Interviews wurde deutlich, dass eine Gemeinsamkeit von ForscherInnen und LehrerInnen, die sich im Young Science-Netzwerk einbringen, darin besteht, dass der Erstkontakt zu Young Science überwiegend projektspezifisch entstanden ist. Häufig stand ein Sparkling Science-Projekt am Beginn der Zusammenarbeit. In manchen Fällen kam der Erstkontakt auch über Young Science-Projekte zum Thema Nachhaltigkeit zu Stande.

Als Barriere für die Vernetzung verschiedener Akteure in Schulen und wissenschaftlichen Einrichtungen erwiesen sich zeitliche Ressourcen und institutionelle Rahmenbedingungen: Von LehrerInnen wurden unflexible schulische Strukturen als eine Herausforderung beschrieben. Nur in wenigen Fällen wurde berichtet, dass Projekte in den Regelunterricht integriert werden konnten. Bei der Teilnahme an Veranstaltungen sind Ansuchen um Freistellung und Abgeltung von Reisekosten notwendig, usw. Dadurch hängt das Gelingen einer Zusammenarbeit stark vom persönlichen Engagement motivierter Lehrkräfte ab, das häufig in der Freizeit stattfindet.

Zeitliche Ressourcen erwiesen sich auch bei der Vernetzung mit ForscherInnen ausschlaggebend. Die Evaluierungsergebnisse verdeut-

lichen, dass der Fokus von WissenschaftlerInnen stark auf der (Grundlagen-)Forschung liegt. Aktivitäten für SchülerInnen werden häufig von engagierten ForscherInnen in ihrer Freizeit organisiert (beispielsweise Führungen oder Exkursionen). Wenngleich zur Unterstützung der Vernetzung von Schulen und Universitäten im Zuge der Leistungsvereinbarung Forschungs-Bildungs-KoordinatorInnen (Kontaktpersonen von Young Science) an österreichischen Universitäten etabliert wurden, verdeutlichen die Evaluierungsergebnisse, dass bei der Etablierung der Kontaktpersonen an Universitäten häufig nicht klar war, welche Erwartungen mit dieser Rolle verbunden sind. Befragte Kontaktpersonen gaben an, dass ihre Rolle als Kontaktperson für Young Science und für Schulen nur einer ihrer Aufgabenbereiche ist und die zeitlichen Ressourcen für diese Aufgaben auch bei ihnen stark beschränkt sind.

Diese Rahmenbedingungen beeinflussen stark die Art und Weise, in der Young Science Informationen bereit stellen muss, damit diese von den verschiedenen Akteuren auch genutzt werden. Das Feedback im Zuge der Evaluierung verdeutlicht, dass Young Science dieser Herausforderung bislang sehr erfolgreich begegnet. Von den verschiedenen Services, die von Young Science angeboten werden, wird vor allem der persönliche Kontakt sehr geschätzt. Bei Fragen oder Problemen einfach anrufen oder ein E-Mail schreiben zu können, erspare Zeit und wird der Recherche auf der Homepage häufig vorgezogen. Der persönliche Kontakt mit MitarbeiterInnen von Young Science wird als unkompliziert wahrgenommen und die prompten Antworten auf Anfragen wurden in Workshops und Interviews wiederholt als sehr positiv hervorgehoben.

AKTIVES ZUGEHEN AUF (NEUE) AKTEURE ERWEIST SICH ALS ERFOLGSREZEPT

Die sehr hohe Flexibilität und Serviceorientierung der Young Science MitarbeiterInnen zeigt sich auch darin, wie Young Science ständig bemüht ist, neue Formate zu entwickeln, um Informationen an die Zielgruppen zu bringen bzw. versucht, im Gedächtnis zu bleiben. Im Rahmen der „Infotour-Veranstaltungen“ etwa kommt „der Prophet zum Berg“ und informiert im Rahmen anderer Veranstaltungen - wie beispielsweise DirektorInnentagungen oder Veranstaltungen für SchülerInnen- und BildungsberaterInnen - über die neuesten Aktivitäten von Young Science.

Dieses ständigen Bemühen von Young Science, seine Netzwerkkontakte aktiv zu halten, wird von den Akteuren im Young Science-Netzwerk sehr positiv bewertet. Workshops und Interviews verdeutlichen, dass besonders der Aspekt, bei einem besuchten Event auch gleich die Neuigkeiten von Young Science mit geliefert zu bekommen (bspw. bei Veranstaltungen die Young Science bei der Infotour besucht), geschätzt wird und diese Form der Informationsbereitstellung erfolgreich macht.

JE ZIELGRUPPENSPEZIFISCHER DIE VERANSTALTUNG DESTO GRÖßER DER NUTZEN FÜR DIE TEILNEHMERINNEN

In Bezug auf Veranstaltungen, die von Young Science selbst organisiert werden, zeigen die Evaluierungsergebnisse, dass diese vor allem dann erfolgreich sind, wenn sie einen klaren praktischen Nutzen für die TeilnehmerInnen haben. Ein gutes Beispiel dafür sind die von Young Science initiierte Vernetzungstreffen der Kinderuniversitäten, die sich an OrganisatorInnen von Kinderuniversitäten richten. Bei diesen AkteurInnen bestand ein klarer Bedarf an Austausch, aber es fehlte ein Mittler, der den Austausch organisiert und strukturiert. Die Personen, die Kinderuniversitäten organisieren, hatten sich zuvor vielfach gegenseitig nicht gekannt.

Die Vernetzungstreffen haben zur Weiterentwicklung der Kinderuniversitäten beigetragen, indem neue Kooperationen entstanden. Auch gemeinsame Projekte wurden bereits angedacht.

ONLINE-SERVICES BÜNDELN INFORMATIONEN ZU AKTIVITÄTEN AN DER SCHNITTSTELLE SCHULE - WISSENSCHAFT

Ein Projektziel von Young Science ist es, Informationen zu Bündeln und Synergien zwischen Initiativen zu nutzen. Dieses Ziel wird durch die Young Science-Website und den Newsletter zu erreichen versucht. Auf der Young Science-Webseite werden Informationen für verschiedene Zielgruppen aufbereitet. Die Webseite stellt einerseits eigene Projekte und Initiativen (wie das Young Science-Gütesiegel und die Themenplattform für vorwissenschaftliche Arbeiten) vor, andererseits bewirbt sie fremde Initiativen, wie Sparkling Science, das Projekt „Studieren probieren“ oder die MaturantInnenberatung.

Zudem kündigt der Young Science-Newsletter ca. einmal im Monat Veranstaltungen, Ausschreibungen und Termine an, die für Personen mit Interesse an der Forschungs-Bildungs-Kooperation relevant sind. Auch im Newsletter werden nicht nur Aktivitäten und Inhalte von Young Science selbst promotet, sondern auch auf andere relevante Veranstaltungen an der Schnittstelle zwischen Bildung und Wissenschaft hingewiesen.

In den letzten Jahren konnte die Webseite steigende Zugriffszahlen und längere durchschnittliche Verweildauern verzeichnen. Auch jährliche Neuanmeldungen zum Newsletter weisen darauf hin, dass das Ziel, Informationen zu bündeln und weiter zu geben, erreicht wird und Bedarf an einer Servicestelle wie Young Science, die Informationen zu einer Vielzahl von Aktivitäten verschiedener Akteure zusammenführt, besteht.

DIE THEMENPLATTFORM UNTERSTÜTZT SCHÜLERINNEN BEI DER THEMENFINDUNG IHRER VORWISSENSCHAFTLICHEN ARBEITEN

Das Verfassen einer vorwissenschaftlichen Abschlussarbeit (VwA) bzw. Diplomarbeit in Schulen, die zur Matura führen (AHS und BHS), ist seit 2012 in Österreich verpflichtend und wird seit dem Schuljahr 2014/15 erstmals flächendeckend eingeführt. Young Science wurde vom BMW-FW damit beauftragt, eine Plattform zu entwickeln, die SchülerInnen bei der Themenfindung für die vorwissenschaftliche Abschlussarbeit unterstützt. Die Themenplattform von Young Science ist eine Online-Plattform auf der Young Science-Webseite, auf der aktuelle Forschungsthemen als Themenanregungen für SchülerInnen aufgelistet sind. Die Themenplattform wird von WissenschaftlerInnen mit Themenvorschlägen befüllt und von SchülerInnen und LehrerInnen genutzt. Die Themenplattform richtet sich daher an drei wesentliche Zielgruppen: SchülerInnen, LehrerInnen und WissenschaftlerInnen.

Im Zuge der Evaluierung wurde deutlich, dass die Einrichtung der Themenplattform von allen Zielgruppen sehr geschätzt wird. Die gesetzliche Einführung der VwA sorgte laut Interviews und Workshops sowohl bei LehrerInnen, SchülerInnen als auch WissenschaftlerInnen³ für große Unklarheiten und wurde besonders von LehrerInnen als große Herausforderung bzw. „akute Baustelle“ (Zitat Lehrerin) empfunden. Die Evaluierungsergebnisse zeigen, dass die auf der Plattform gebündelten Informationen rund um die vorwissenschaftlichen Arbeiten und Verlinkungen zu anderen relevanten Initiativen, wie z.B. Infowebseite des Ministeriums für Bildung und Frauen (BMBF) zu VwAs, manche Unklarheiten bei LehrerInnen, SchülerInnen als auch bei WissenschaftlerInnen beseitigen. Die Vielfalt an Themenvorschlägen auf der Plattform wird nicht nur von SchülerInnen sondern auch von LehrerInnen geschätzt. LehrerInnen betonten, dass sie seit Einführung der verpflichtenden VwA verstärkt mit Themenanfragen von SchülerInnen konfrontiert sind, die oft ihren eigenen Themenbereich übersteigen. Die Themenplattform bietet für sie eine Unterstützung, um SchülerInnen dennoch kompetent beraten zu können.

ABSTIMMUNGSPROBLEME UND ABGRENZUNGSSCHWIERIGKEITEN BEI DER ENTWICKLUNG DER THEMENPLATTFORM

Im Zuge der Evaluierung wurden auch Schwierigkeiten bei der Entwicklung der Themenplattform sichtbar, die für die Arbeit an Schnittstellen typisch sind. Diese Schwierigkeiten betrafen Abstimmungsprobleme mit anderen Institutionen, sowie Abgrenzungsschwierigkeiten in Bezug auf die Aufgabengebiete und die Rolle von Young Science gegenüber den verschiedenen Zielgruppen.

Neben Unklarheiten bezüglich der generellen Anforderungen einer VwA, kam es durch Aktivitäten verschiedener Organisationen, wie etwa Initiativen des Bundesministeriums für Bildung und Frauen (BMBF) sowie von Young Science zu Mehrfachanfragen an ForscherInnen. Wenn gleich es sich teilweise um Anfragen unterschiedlichen Inhalts handelte, war dies für ForscherInnen nicht immer klar ersichtlich. In den Interviews und Workshops im Rahmen der Evaluierung wurde wiederholt betont, dass dies für Unmut bei den Betroffenen sorgte.

Die Evaluierung machte zudem viele Unklarheiten von Seiten der LehrerInnen als auch der WissenschaftlerInnen bezüglich der relativ neuen gesetzlichen Regelung deutlich. Durch den hohen Bedarf an Information und Unterstützung in Zusammenhang mit der VwA, der nicht nur an Universitäten sondern auch an Schulen besteht, wurden in den Interviews und Workshops Erwartungen an Young Science artikuliert (wie etwa Hilfestellung für Lehrpersonen in Zusammenhang mit sozialwissenschaftlicher Methodik), die laut MitarbeiterInnen von Young Science außerhalb ihrer Zuständigkeiten liegen. Das unterstreicht die Bedeutung von guter Abstimmung und klarer Abgrenzung von Zuständigkeiten bei der Arbeit an der Schnittstelle zwischen Schule und Wissenschaft.

MIT DEM YOUNG-SCIENCE GÜTESIEGEL WIRD EIN ANREIZ FÜR FORSCHUNGS-BILDUNGS-KOOPERATION GESCHAFFEN

Das Young Science-Gütesiegel zielt darauf ab, einen Anreiz für Forschungs-Bildungs-Kooperation zu schaffen, indem besonders engagierte Schulen ausgezeichnet werden. Das Gütesiegel soll der Forschungs-Bildungs-Kooperation auch zu mehr Sichtbarkeit verhelfen.

Seit 2012 zeichnet das Gütesiegel Schulen mit einem herausragenden Engagement in der Forschungs-Bildungs-Kooperation aus. Die Bewerbung für die Auszeichnung steht allen Schulen und Schultypen offen. Bislang wurden insgesamt 33 Schulen mit dem Gütesiegel ausgezeichnet. Die Evaluierung zeigt, dass die mit dem Gütesiegel ausgezeichnete Schulen sowohl intern (von Young Science selbst) als auch extern

(von Universitäten) als mögliche Partnerschulen für Projekte herangezogen werden. Eine Broschüre mit den ausgezeichneten Schulen wird von Young Science bei Treffen mit Hochschulen benutzt, um Schulen *„ohne schlechtes Gewissen weiterzuempfehlen“* (Zitat Young Science Mitarbeiterin).

Die Ergebnisse der Evaluierung machen deutlich, dass das Young Science-Gütesiegel für einige Kontaktstellen an Universitäten ein wichtiges Qualitätsmerkmal ist, wenn Schulen als Kooperationspartner für Projekte gesucht werden. Andere wissenschaftliche Akteure betonen hingegen, dass sie Schulen aufgrund persönlicher Kontakte auswählen.

Aber auch LehrerInnen betonten den Nutzen des Gütesiegels für Schulen. Die Auszeichnung erhöhe die Wettbewerbsfähigkeit von Schulen, indem die Schulen das Gütesiegel beim Werben um neue SchülerInnen als „Aushängeschild“ verwenden.

Somit wird mit dem Young Science Gütesiegel einerseits die Sichtbarkeit von Schulen gestärkt, die erfolgreich mit Forschungseinrichtungen kooperieren, und ein Anreiz für die Forschungs-Bildungs-Kooperation geschaffen. Andererseits fällt es leicht, ausgezeichnete Schulen an Forschungseinrichtungen zu vermitteln, da sichergestellt ist, dass diese Schulen über Erfahrung in der Projektumsetzung verfügen und bereits erfolgreich mit WissenschaftlerInnen zusammengearbeitet haben.

EMPFEHLUNGEN

In Bezug auf die Vernetzungsaktivitäten empfiehlt es sich, „aufsuchende“ Informationsangebote (wie bspw. die „Infotour“) weiter auszubauen. Die persönliche Kontaktebene wurde von den Befragten wiederholt als äußerst wichtig hervorgehoben und sollte somit weiter gestärkt werden.

Eine kürzlich gestartete Initiative von Young Science, ForscherInnen dafür zu gewinnen, als BotschafterInnen an Schulen zu fungieren, kann als Schritt in diese Richtung bewertet werden. Es böte sich auch an, dass Lehrpersonen als BotschafterInnen an anderen Schulen ihre Aktivitäten im Rahmen der Forschungs-Bildungs-Kooperation vorstellen und über das Gütesiegel informieren, um neue Schulen für die Forschungs-Bildungs-Kooperation zu gewinnen.

Um zukünftig auch Pädagogische Hochschulen stärker in das Netzwerk einzubinden, sollten Anstrengungen zur Etablierung von Young Science-Kontaktpersonen an den Pädagogischen Hochschulen forciert werden. Dies würde zu einer stärkeren Öffnung von Young Science gegenüber Schultypen wie Neuen Mittelschulen und Volksschulen beitragen und gleichzeitig könnten angehende LehrerInnen noch während ihrer Ausbildung für die Forschungs-Bildungs-Kooperation interessiert und das Young Science-Netzwerk dadurch erweitert werden.

Auch das Schaffen konkreter Anreize, wie einer Anerkennung des Engagements von LehrerInnen und SchülerInnen (z.B. über Zertifikate), könnte für den Ausbau des Netzwerkes hilfreich sein.

In Hinblick auf Veranstaltungen, die von Young Science selbst organisiert wurden, zeigt sich, dass Veranstaltungen vor allem dann von den TeilnehmerInnen als wertvoll wahrgenommen wurden, wenn sie einen

klaren Nutzen bzw. Mehrwert für diese hatten (z.B. Vernetzungstreffen der Kinderuniversitäten). Daher sollten kleinere Formate, die sich mit einem konkreten Thema beschäftigen und sich an eine spezifische Zielgruppe richten, verstärkt angeboten werden.

Sowohl für WissenschaftlerInnen als auch LehrerInnen könnten thematische Austauschtreffen, bspw. zu Herausforderungen der Projektumsetzung, einen großen Mehrwert darstellen. Diese sollten in kleinem Rahmen (z.B. Workshop) und auch in den Bundesländern stattfinden.

Bei der Themenplattform verdeutlichen sich Schwierigkeiten, die an Schnittstellen wie der zwischen Schule und Wissenschaft immer wieder auftreten - durch Zuständigkeiten verschiedener Institutionen können Doppelgleisigkeiten entstehen, die bei Personen, die wenig zeitliche Ressourcen zur Verfügung haben, für Unmut sorgen. Hier scheint eine gute Abstimmung bzw. eine Kooperation der involvierten Organisationen essentiell. Eine solche Kooperation sollte in Zukunft noch stärker ausgebaut werden. Zudem empfiehlt sich eine klare Abgrenzung der eigenen Zuständigkeiten zu anderen Initiativen und eine verstärkte Kommunikation über Rollen und Aufgabengebiete von Young Science in Bezug auf die VwA, damit keine Erwartungen entstehen, die nicht erfüllt werden können.

AUTORINNEN

CAROLINE MANAHL

Zentrum für Soziale Innovation

E: manahl@zsi.at

SUSANNE DOBNER

Zentrum für Soziale Innovation

E: dobner@zsi.at

WHEN TWO INTERMEDIARIES AND ONE FUNDING MINISTRY MEET FOR A COMMON GOAL: THE CASE OF THE FEDERAL AUSTRIAN INITIATIVE “EVOLVE” FOR THE CREATIVE INDUSTRIES

ALFRED RADAUER AND TOBIAS DUDENBOSTEL

INTRODUCTION

It is only some ten years ago, when innovation and research policy in Austria discovered the topic of ‘Creative Industries’ (CI) as a worthwhile field of activity. Spurred by an initiative of the Austrian Federal Chamber of Commerce (WKO), and implemented as a joint effort with the Austrian funding bank aws (“Austria Wirtschaftsservice GmbH”), the first attempts to support the Creative Industries included the establishment of a grant scheme (“Impulse Programme Creative Industries”) and some soft support measures in the area of training and awareness raising.¹ The challenges noted then were manifold: While the Creative Industries were understood to be an important driver for innovation in itself, and also supplying services relevant to innovation in non-creative sectors, it was difficult to pin down who (in terms of what firms, actors) made up the sector or how innovation processes functioned in detail in that sector.

The firm structure in the innovative creative industries – somewhere located in-between the cultural sector and the technology-oriented innovating business sectors – was mostly composed of very small firms that had their own culture. For example, the firms neither considered themselves as businesses (rather as “artists”), nor were they aware that that they belonged to something called “creative industries”. One core issue in this context was to find an appropriate workable definition of the CI – an issue, which has to date not found consensus across policy makers in different countries as well as among the international cast of scholars in this field.² According to the current definition of the CI used in Austria, the CI were made up, in 2012, of some 39,000 companies (which corresponded to some 10.4% of all Austrian firms) with around 140,000 employees. Overall industry turnover was € 20 billion.³

Fostering the CI in adequate manner proved to be difficult also from an institutional point of view. The two involved organisations were faced with the challenge to create the grounds for support activity where little had existed before specifically for the creative field. On the one hand, the working group for the CI at the Austrian Federal Chamber of Commerce (“creativwirtschaft austria” or cwa) was caught in-between the very traditional sector-oriented organisation of the Austrian Chamber of Commerce, and the creative industries firms that displayed the said “anti-business” attitude and had a rather reluctant stance towards the Chamber of Commerce overall. aws, in turn, as funding bank for industry and business overall had to deal with this “distance” from the Creative Industries firms, too. Furthermore, internally there was also a need to sell and establish CI support against proven support programmes and funding schemes in technological areas. Eventually, there was also a need to collaborate between aws and cwa, with all opportunities and challenges that may arise in such inter-agency collaborations.

Since the times of the “impulse programme”, the respective support measures for the creative industries were evaluated and monitored closely and improved. Intelligence was gathered – e.g. in the form of the biannual reports on the state of the Creative Industries –, and a new funding-providing body for the grants and activities (the Austrian Federal Ministry of Economy, Family and Youth (BMWFJ)) replaced the previous funder. By 2009, the different offers of the aws and the cwa for the “Creative Industries” were re-organised under the umbrella initiative “evolve”. The set-up was as follows: The Federal Ministry of Economy, Family and Youth (BMWFJ) was the programme owner and was funding the initiative. The Ministry acted at the strategic level and defined the objectives. Operational programme management was handled by said

1 see, for example, Radauer, A. & Dörflinger, A. (2007): Begleitende Evaluierung des „iP ImpulsProgramm creativwirtschaft“

2 It would go beyond the scope of this paper to describe the issues at hand in greater detail. However, one major issue worth mentioning is that many definitions may be too broad for practical purposes like the design of a support programme. For a more comprehensive discussion of different approaches to define the CI, see, for example, UNCTAD (2008): Creative Economy Report 2008.

3 Gassler et al. (2015): Sechster österreichischer Kreativwirtschaftsbericht (“Sixth Report on the Austrian Creative Industries”). The Austrian CI definition is NACE-based and comprises nine sectors: architecture; design; music, book and artistic activity; radio and TV; software and games; publishing; video and film; advertisement; libraries, museums and botanic and zoological gardens.

two organisations: on the one hand the Austrian Federal funding bank aws, and on the other hand the cwa. Both organisations implemented different “evolve” measures.

By 2013, it was time to look at the achievements of the initiative, how cwa and aws had managed the challenges and what learning effects were to be recorded for the future initiative “evolve 2.0”, also and particularly with respect to the viability of the chosen institutional set-up. The respective evaluation was contracted by the BMWFJ to Technopolis Group Austria. Methodically, the evaluation was based on a mix of qualitative and quantitative instruments: the analyses of existing studies and monitoring data, a standardised survey on recipients of project funding of the instruments “impulse XS” and “impulse XL” (with 983 sent-out questionnaires, 186 returns and a satisfying return rate of 18.9%), two focus groups with CI experts, one international workshop, 29 qualitative interviews, three case studies on funded projects and a participating observation in the course of a jury meeting.

Because of the requirements of the BMWFJ and the already well-defined existing data and information from previous monitoring and evaluations, there was a need to implement a methodology that allowed us to drill deeper into the subject matter. An analysis of funded projects at micro level (i.e., an analysis of project contents) determined who or what exactly was funded by “evolve” beyond what could be gathered from the analysis of mere monitoring data. This latter scrutiny led to the selection of the said three case studies in order to illustrate typically funded projects. The inclusion of three international experts for the creative industries (from Denmark, England and Germany) allowed the evaluation team also to consider international experiences in detail.

FINDINGS

The following findings emerged:

- „evolve” can be characterised by three pillars that are to lead to the achievement of the overarching objectives of the initiative: “to harvest the innovative potential of the CI” and “to increase the competitive capacity in the area of innovation”. The first pillar - and the financially most important one - is monetary funding for innovative projects. It consists of the three funding instruments “impulse XS”, “impulse XL” and “impulse LEAD”. The second pillar, concerned with services, covers educational offers, networking activities and further services. The third pillar, dealing with awareness raising, focuses on raising visibility of the Creative Industries and their potential as an industry, as well as advertising the funding and services “evolve” offers. Depending on the extent that the various “evolve” activities can be regarded as measures in their own right, “evolve” consists of 18 or more single measures with a variety of target groups.
- The overall result of the evaluation was positive. Most of the various “evolve” measures were used quite intensively and performed well. Coordination between the programme managing organisations aws and cwa was satisfactory. Concerning the internal administration of the various measures, the activities, instruments and objectives in aws and cws were clearly structured, well delimited and connected in a logical way. Via the initiative “evolve”, the preceding programmes and offerings of the cwa of the WKO and the aws were developed further in a meaningful way. Possible areas of friction between the aws and the cwa were minimised.
- aws and cwa managed their activities and instruments professionally. Users, funding beneficiaries and stakeholders alike emphasized the management teams’ efforts and competence, also by comparison to other Austrian RTDI initiatives.
- The evaluation team’s inquiries showed that the initiative led to a broad spectrum of impacts. Both for “impulse XS” and in “impulse XL”, funding beneficiaries reported a high benefit of the funding, especially as a contribution to the development of the company resp. the business model, for helping to secure and raise turnover, for securing and boosting employment, for accessing new sources of finance or for building reputation. The additionality of the funding instruments “impulse XS” and “impulse XL” was determined to be high.
- A central milestone for “evolve”, specifically for “impulse XS” and “impulse XL”, was the development of a common understanding and interpretation of what innovation means in the context of the Creative Industries. Such an understanding was necessary because the initiative uses funds that are earmarked for RTDI projects, which means that the “evolve”-specific interpretation of innovation must be compatible with the notion of innovation in RTDI support. Following the observation that innovation in the CI is not so much about research, one main challenge was to take experimental development, as it is regularly happening in the CI, in a sensible way into consideration for defining the funding criteria of “impulse XS” and “impulse XL” projects.
- Within “impulse”, innovation was defined as a process with experimental development character necessary for solving problems using CI-based outputs. The usage of the notion of experimental development for problem solving is particularly important in order to discern “impulse” funding from pure arts and culture funding. A new design for a chair, for example, does per se not fulfil the funding criteria for “impulse”. However, if the design of the chair solves also a specific problem – such as to help handicapped persons to stand up/sit down – it would be an innovation eligible for funding, provided that there is also some experimental development work involved. This notion of innovation is also applicable to service innovations. For example, one project funded under the title of URBANAUTS used architecture and design in combination with a new business model to convert the empty space of former street shops into decentralised and autonomously operating hotel rooms. The experimental development involved included technical solutions to improve safety, noise insulation or new booking systems that would allow travellers to choose their rooms individually (rather than, with a traditional hotel, book one type of room and have the reception assign a room number upon arrival). Further examples covered

the innovative media area and frequently involve the creation of software solutions as part of the project activities.

- The whole process of project selection is depicted in the figure below. It is a three-stage jury-based process: In the first stage, “impulse” looked for projects that can be submitted by SMEs (in the case of “impulse XL”) or by micro-enterprises/physical persons (in the case of “impulse XS”). While the submitting firms and/or physical persons could be from any industry, the projects must have been in or be related to one of ten listed domains of the CI (see figure below). These projects must, secondly, have worked on innovations and involved experimental

development work as described above. In the third stage, the projects were checked as to whether they support the existing or an envisaged new business model of the submitting firms/ persons. Because “impulse XS” and “impulse XL” targeted different maturity levels – “impulse XS” being concerned with the proof of feasibility of an innovation, while “impulse XL” aiming for the innovations themselves – the criterion of supporting a viable business model is treated differently in “impulse XS” and “impulse XL”. The result of the funding process was either a proven feasibility of an innovation (“impulse XS”) or an innovation as such (“impulse XL”).

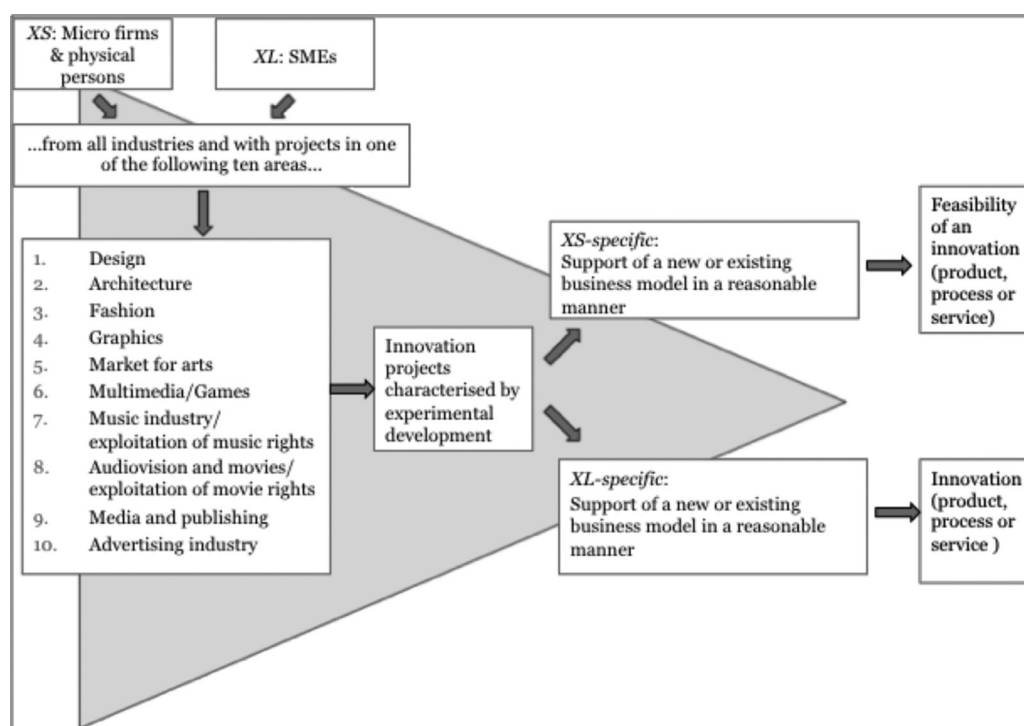


Figure1: Selection process for projects funded under „impulse XS“ and „impulse XL“

Source: Interviews und documents from „impulse“, own illustration by Technopolis Group Austria

- The respective efforts for finding a specific interpretation of innovation resulted in the initiative’s ability to fund particular types of innovation projects that have considerably lower (and/or non-existing) technical/engineering requirements than “traditional” RTDI schemes. The specific interpretation of “innovation” was a rather broad one. While it took a considerable amount of time to develop and operationalize the specific notion of innovation in the context of “evolve”, the compiled evidence in the evaluation suggests that the former challenge to deal with innovation and experimental development in a CI-funding context had turned into an asset of the initiative.
- In particular, the broad understanding of innovation led, in combination with the jury selection process (that makes use of national and international jurors), to the possibility to indirectly fund specific business model innovations. These are business

models that are based on the CI-outputs and that back an innovation generated by an experimental development process. This is remarkable because traditional RTDI funding programmes find it for the most part difficult to deal with the topic of business model innovation, even though the importance of business model innovation for transferring R&D results to the market has been increasingly recognised. The jury selection process required the applying firms or physical persons to explain the business case for the project. Funding was only granted when the respective explanation was convincing. Hence, “impulse” funding may lead to situations where different forms of business model innovation were considered in the form of a precondition for the economic success of the project.

- Nonetheless, advantages and possibilities of this asset of a specific notion and interpretation of innovation were not used to

their full potential within “evolve”. Because it was difficult to explain in a straightforward manner the specific interpretation of “innovation” to the target audiences, “impulse” used case studies and examples as well as 1:1 advice to some effect to this end – which is a sensitive approach. However, seemingly unsatisfying explanations of why some proposals had been rejected seem to have undermined those efforts to some extent. Insufficient explanations led to confusion in regard to the understanding of what is funded and, in succession, to negative ratings of aspects like “feedback to refused proposals” and “transparency of the decision making process” in the evaluation’s survey of applicants for the two funding schemes. The problem was exacerbated by the fact that demand for funding was quite high, leading to low success rates for proposals and hence also to some good proposals not obtaining funding.

- The instrument „impulse LEAD” was a measure directed at collaborative networks with involvement from SMEs that implement projects with role model character for the future and which aim at networking, professionalising or increasing the visibility of the CI. The compiled evidence suggests that “impulse LEAD” suffered from an overload of objectives. Therefore, various interview partners doubted its overall usefulness and effectiveness. At the same time, our interviewees repeatedly emphasized the necessity and benefit of an instrument like “impulse LEAD” to create structures for the CI or as a tool for “institutional learning”. Despite of the generally mediocre feedback on “impulse LEAD”, there were a number of well-performing funded projects that also obtained international attention and acclaim. The evaluation team developed several options on how to develop “impulse LEAD” further.
- Another evaluation result was the high benefit and usefulness of the various complementary awareness raising measures and services, for example in the area of training and education or networking. The demand for those offers was high, the satisfaction with them was on average very high and the expectations, for example with regard to the establishment of communities, were for the most part met. The complementary offerings were an important link between the administering/managing organisations aws and cwa and their target groups. They gave the management teams a better knowledge about their customers. However, some interviewees complained about a missing overall concept explaining what offerings can be combined with others in a meaningful way particularly in the area of training.
- Over time, “evolve” enhanced the links between the CI and other sectors of the economy to some, albeit rather small, extent. This was a specific question for the evaluation, as it is understood that the CI have also an “enabling” function for innovations in more “traditional” sectors of the economy. However, linking the CI with other economic sectors remained an important and largely untapped area of activity for “evolve” mainly due to the reason that there were only few offerings that explicitly focussed on this type of link. More precisely, while there were measures that could (and sometimes are) used to link the CI with other industries, the respective possibilities were hardly marketed. An important step for fostering the CI/non-CI link was the then-recent creation of a voucher scheme for the Creative Industries, whereby firms in other economic sectors could make use of CI firm services. This scheme was in high demand, but was – due to its then young age – not subject of this evaluation.
- Regarding the possibility to foster the links between the CI and other economic sectors, many of the interviewees expressed considerable expectations that the cwa would be in a position to tackle this issue more profoundly. The reason is that because the cwa is part of the Austrian Federal Economic Chamber, it was seen to have “privileged” access to such “traditional” firms that are members of the chamber and to be therefore in a position to easily bridge the gap between the CI and the non-CI industries. It has to be remembered that Austria has compulsory membership for all firms in the chamber of commerce.
- However, there were a number of factors that stood against the wider realisation of the said expectations. First, the cwa was tasked specifically with the execution of a range of specified services and awareness raising activities that did not have, for the most part, a specific focus on bridging the gap between CI and other industries. Moreover, because of the strategic choice to focus on service execution and awareness raising, there was little activity and budget left in the outside communication/marketing for clarifying to the target groups in a more profound manner what the cwa as an institution did and could do within its budget limits. For example, the cwa did not manage project funding (what is the task of the aws) which limited its ability to initiate cross-sectorial projects.
- At the level of single measures and instruments, most activities managed by the cwa were rated nonetheless positively and were in high demand. The cwa managed to create a portfolio of activities that was an important factor for the cwa’s acceptance in the CI, and it also fulfilled hereby a bridge for CI firms that were traditionally sceptical towards a business chamber.
- The evaluation team found that “evolve” was not a widely known brand among the beneficiaries of the various support measures of the initiative. Beneficiaries knew the individual measures, but they did not know the broader concept behind the measures and how the managing organisations, the funding ministry and the respective activities interrelated to and with each other. Apart from the fact that the lack of knowledge of the “evolve” brand may have contributed to a small extent to the rather unclear picture of the institutional role of the cwa, the evaluation team did not see the weak perception of the “evolve”-brand to have been too much of an issue. After all, beneficiaries do only need to know and use the services and funding schemes that are being offered to them in the desired way. They do not need to know the mechanics behind these measures. As a result, the evaluation team saw little need to invest in branding “evolve” more. To the inside – that is, among the stakeholders of the initiative – the brand fulfilled nonetheless an important coordinating function.

CONCLUSIONS

The evaluation of “evolve” aimed to analyse the initiative both in terms of its past performance (ex-post aspect) as well as in terms of its future potential (ex-ante aspect). It showed that, over the years – and particularly due to highly committed staff at both cwa and aws – it was possible to create a well-functioning machinery that supported the Austrian Creative Industries. The chosen institutional set-up allowed for good working division and minimised inter-agency friction. Challenges in reaching out to the target group of “Creative Industries” firms were mostly mastered. However, there were also shortcomings: e.g., the difficulties to find answers on how to support cross-sectorial innovation projects; or the problems identified with “impulse LEAD”.

The joint BMWFJ/aws/cwa concept for the then-upcoming and updated “evolve 2.0” initiative – at that time available only in draft form – mainly aimed to continue the previous measures and activities of the evaluated “evolve” initiative. For those measures where problems had been already acknowledged (such as with “impulse LEAD”), the concept called for a revision. There were as well a few new measures, but the basic concept behind “evolve” remained the same. In the end, the concept “evolve 2.0” was the result of positive experiences with the then-current “evolve” initiative, whereby “evolve” was understood to be for the most part a means to coordinate the activities of the two organisations aws and cwa and the BMWFJ in an effective and efficient manner. A possible danger of this approach – as seen by the evaluation team – was a potential lock-in situation, meaning that because of the success of the existing structures, completely new or different approaches and opportunities to fostering the CI may not be recognised.

“Traditional” RTDI policy making can potentially learn from two features of “evolve”. The developed broad notion of innovation as a process with experimental development character using the outputs of the CI is to be noted first. It allowed “evolve” to support non-technological innovations (or, to be more precise, less technologically geared projects where the technological content would not be usually funded in a competitive RTDI grant programme). Moreover, it supported – indirectly – certain forms of business model innovations. The second issue concerned how the two involved agencies interacted. The general model of separating monetary support and making it a task of one agency, whereas the other agency is given responsibility for complementary “soft” measures, data collection and awareness raising – and this set-up being governed by an umbrella initiative, where all involved agencies interact with each other intensively – could prove to be a viable approach in two-agency institutional set-ups working jointly on the same areas of support. The one challenge of “evolve” already noted – the difficulty in supporting cross-sectorial innovation – could also be seen as starting point to bring “traditional” RTDI support and CI support institutions closer together.

The evaluation team spelled out a number of recommendations. At the strategic level the team recommended:

- That “evolve 2.0” be transformed into a wider-reaching/broader strategy of the BMWFJ for innovation in the CI, avoiding thus the issue of a possible “lock-in” and giving the initiative an enhanced ability for new approaches to foster the CI

- That specific offerings for bridging the gap between the CI and other industries be implemented in “evolve” to a larger extent and communicated in a reasonable manner to the target audiences
- That the role of the cwa as institution, after an examination of the feasibility and a check whether this is in line with future strategic considerations, be sharpened and better communicated/marketted
- That the new and low-threshold instrument of Creative Industries Vouchers should be evaluated

At the level of single measures or instruments the team recommended:

- That most of the „evolve” measures should be continued, but only after strategic considerations and prioritising, which can very well lead to the cancellation of single and well working instruments
- That the design of „impulse LEAD” should be considerably improved for which there are a variety of options (one of the options would also be to cancel the instrument)
- To examine the necessity for the instrument “CreativDepot” – a measure of the cwa that allows the registration of copyrights for proving the date of creation (“time stamp service”) – and if found to be in line with future strategic considerations, to make it part of existing extensive broad IPR service portfolios of specialised agencies and organisations in this area, e.g. that of the aws in the area of IPR (the aws had its own department on IP and licensing)
- That the instruments “impulse XS” and “impulse XL” should be continued with slight adaptations
- That a stronger common structure covering the set of training and awareness measures should be created and communicated accordingly

Eventually, a suggestion for a results/impact-based monitoring system was developed, too, involving the set-up of a combined qualitative and quantitative assessment grid.

AUTHORS

ALFRED RADAUER

Technopolis Group Austria

E: alfred.radauer@technopolis-group.com

TOBIAS DUDENBOSTEL

Technopolis Group Austria

E: tobias.dudenbostel@technopolis-group.com

AUSWAHLVERFAHREN VON FORSCHUNGSFÖRDERUNGSORGANISATIONEN IM INTERNATIONALEN VERGLEICH

PETER BIEGELBAUER UND THOMAS PALFINGER

Der Erfolg eines Innovationssystems hängt von der Qualität jener Tätigkeiten ab, die den Innovationen vorgelagert sind und zu eben jenen führen sollen. In wesentlicher Art und Weise sind das Forschungstätigkeiten, die üblicherweise im Rahmen von Projekten durchgeführt werden (Dinges 2010). Aufgrund der Annahme von Marktversagen (Arrow 1962, Biegelbauer 2000, Kubeczko/Weber 2009) werden Auswahl, Finanzierung, Begleitung und Evaluation derartiger Projekte heute zumeist von Forschungsförderungsagenturen mit öffentlichen Geldern bestritten (Braun 1997, Christensen/Laegreid 2006, Verhoest et al 2007, Weber 2009).

Dadurch erhalten jene Auswahlverfahren besondere Bedeutung, auf deren Basis die zu fördernden Projekte ausgesucht werden. Gelingt es nicht in ausreichendem Maße jene Projekte, die für die Entwicklung eines Innovationssystems nützlich sind, von denjenigen, die das weniger sind, zu unterscheiden, werden die Ziele von Förderprogrammen nicht erreicht und potenziell öffentliche Gelder fehl-alloziert (Donovan 2007, Lepori et al 2007).

Dieser Beitrag beruht auf einer vom österreichischen Bundesministerium für Verkehr, Innovation und Technologie (BMVIT) in Auftrag gegebenen Studie (vergleiche auch Biegelbauer/Palfinger 2014). Er beschäftigt sich mit denjenigen Auswahlverfahren, die von Forschungsförderungsagenturen eingesetzt werden, um Projektanträge effizient auszuwählen, die eine möglichst effektive Umsetzung der Zielsetzungen der Förderprogramme erlauben. Zu diesem Zweck wurden neun Forschungsförderungsagenturen in Ländern ausgewählt, die zu den EU Innovation Leaders und Followern (EC 2013, 2015) gehören.

Dabei handelt es sich um den Innovation Fund Denmark (IFD), den Research Council of Norway (RCN) in Norwegen, die Verket för innovationssystem (Vinnova) in Schweden, Teknologian kehittämisskeskus (Tekes) in Finnland, die Forschungsförderungsgesellschaft (FFG) in Österreich, die Technologické agentury ČR (TACR) in Tschechien sowie die deutschen Projektträger Jülich (PTJ), Technischer Überwachungsverein Rheinland (PT TÜV Rheinland), Verein Deutscher Ingenieure-Technologiezentrum (PT VDI-TZ) und Verein Deutscher Ingenieure/Verein Deutscher Elektrotechniker-Innovation und Technik (PT VDI/VDE-IT).

Der Artikel ist wie folgt gegliedert: Zuerst werden Methodologie und Forschungsfragen geklärt. Dann werden die zehn Forschungsförderungsorganisationen und ihre Projektauswahlverfahren vergleichend

untersucht. Schließlich werden einige Aspekte dieser Auswahlverfahren als modellhafte Praktiken und Entwicklungstrends beleuchtet.

METHODOLOGIE UND FORSCHUNGSFRAGEN

In die Studie eingeflossen sind die Ergebnisse von Literaturrecherchen zu Forschungsförderungsorganisationen, deren organisationale Prozeduren und hier insbesondere Projektauswahlverfahren. Zudem wurden 13 Hintergrundgespräche mit FTI-ExpertInnen, Projekt-GutachterInnen, Programm-EvaluatorInnen und (ehemaligen) MitarbeiterInnen von Forschungsförderungsorganisationen und zehn halbstandardisierte zwischen 45 und 90 min. lange Telefoninterviews mit MitarbeiterInnen von Forschungsförderungsorganisationen geführt.

Die wichtigste Quelle für die Studie waren Telefoninterviews, in denen die meisten Daten zu den Projektauswahlverfahren ermittelt wurden. Die Fragen orientierten sich an den Auswahlverfahren von Forschungsförderungsorganisationen im internationalen Vergleich, insbesondere den Rollen verschiedener Akteursgruppen in diesem Prozess und dem Verhältnis dieser zueinander.

Die leitenden Forschungsfragen waren:

- Sind die Projektauswahlverfahren einheitlich oder nicht?
- Welche Rolle spielen interne und externe GutachterInnen?
- Welche Formen der Projektbegutachtung lassen sich unterscheiden?
- Werden Gutachten kalibriert?
- Wer ist für die Begutachtung und wer für die Finanzierung zuständig?
- Wie wird aus Erfahrung gelernt?

VERGLEICH SPEZIFISCHER ASPEKTE DER PROJEKTAUSWAHLVERFAHREN

Es gibt in allen Agenturen Variationen innerhalb der Auswahlverfahren. Dabei reicht die Bandbreite allerdings von marginalen Unterschie-

den (RCN: im Bereich Innovation sind die meisten GutachterInnen aus dem Bereich der Industrie, im Bereich Grundlagenforschung kommen die GutachterInnen in erster Linie von Universitäten) bis hin zu komplett unterschiedlichen Verfahren innerhalb einer Förderungsstelle (deutsche Projektträger).

Im finnischen Tekes sind die Projektauswahlverfahren im Wesentlichen identisch. Eine der wenigen Unterschiede bezieht sich darauf, dass für Anträge mit Firmen als Leadpartner offene Calls ausgeschrieben werden, im Gegensatz zu Anträgen, bei denen Forschungsorganisationen den Lead übernehmen können, wo es feste Deadlines gibt. Im norwegischen Department Innovation des RCN sind die Unterschiede marginal und beziehen sich vor allem auf unterschiedliche spezielle Kriterien einzelner Programme. Innerhalb der TACR sind die Auswahlverfahren in den unterschiedlichen Programmen sehr ähnlich – auch weil eine Reihe von institutionellen und prozessoralen Elementen im Gesetz, welches der Agentur als Grundlage dient, festgelegt sind.

Im Rahmen eines Reformprozesses in Dänemark werden die Ausschreibungen zu angewandter Forschung seit 1. April 2014 vom IFD ausgeführt. Im Zuge dieser Reform wurden die Auswahlverfahren innerhalb des IFD vereinheitlicht. In der schwedischen Vinnova gibt es eine Reihe unterschiedlicher Verfahren, die in Bezug auf Ein- oder Mehrstufigkeit, dem Einsatz interner oder externer EvaluatorenInnen et cetera auseinander gehen. Allerdings sind einige Auswahlverfahren in den letzten Jahren an die Prozeduren des als besonders erfolgreich wahrgenommenen Programms „Challenge-Driven Innovation“ angenähert worden. So wurden Elemente dieses Programmes in der Gesamtorganisation übernommen.

In der österreichischen FFG sind die Auswahlverfahren in einigen der Bereiche einheitlich, in anderen nicht. Bei den deutschen Projektträgern Jülich, TÜV Rheinland, VDI/VDE-IT und VDI-TZ variieren zwischen, wie auch innerhalb der einzelnen Organisationen die meisten Eigenschaften des Projektauswahlverfahrens.

In Tabelle 1 wird angeführt wie einheitlich beziehungsweise unterschiedlich die Auswahlverfahren in den untersuchten Institutionen sind.

Auswahlverfahren	Fälle
Weitgehend einheitlich	TACR Tekes RCN (seit 2011) IFD
Teilweise einheitlich	Vinnova PT VDI-TZ PTJ FFG
Weitgehend uneinheitlich	PT TÜV PT VDI/VDE-IT

Tabelle 1: Einheitlichkeit der Auswahlverfahren

In Tabelle 2 wird die Häufigkeit des Einsatzes interner bzw. externer GutachterInnen wiedergegeben. Die meisten Forschungsförderungsorganisationen verfügen über eine Mischung der beiden Gutachterttypen, allerdings mit Unterschieden in Bezug auf die Häufigkeit und Bedeutung des Einsatzes der internen und externen GutachterInnen.

Der PTJ ist von einer ausschließlich internen Begutachtung, die vor allem mit Vertraulichkeit begründet wurde, im Jahr 2007/2008 in manchen Forschungsfeldern auf größtenteils externe (nationale) GutachterInnen umgestiegen. Dies geschah vor allem vor dem Hintergrund, dass durch den verstärkten Einsatz von externen GutachterInnen eine bessere Expertise erhofft wurde. Mit Vertraulichkeit begründet auch Tekes sein völlig internes Verfahren. Der PT VDI-TZ hingegen führt die interne Begutachtung im Bereich der Photonik auf einen zu kleinen Kreis an qualifizierten GutachterInnen (hauptsächlich potenzielle MitbewerberInnen) im nationalen und zu hohen (Reise-)Kosten beim Hinzuziehen von ExpertInnen im internationalen Umfeld zurück. Vinnova arbeitete bis 2012 hauptsächlich mit internen GutachterInnen. Dies wurde ebenfalls mit einem Mangel an nationalen GutachterInnen (die Anträge werden auf Schwedisch eingereicht) und zu hohen Kosten (vor allem bei kleineren Programmen) begründet. Im Rahmen der zuvor erwähnten Anpassungsprozesse setzt man seit 2013 wieder verstärkt externe ExpertInnen ein, da man sich dadurch mehr Glaubwürdigkeit bei der Projektevaluation erhofft.

In der FFG werden hauptsächlich externe ExpertInnen eingesetzt. Diese stammen häufig aus einem internationalen (meistens aber deutschsprachigen) Umfeld, um so auch in einem kleinen Land die nötige Unabhängigkeit der GutachterInnen zu sichern. Aus ähnlichen Gründen setzt TACR ausschließlich externe GutachterInnen ein, die in der Regel national rekrutiert werden. Der dänische IFD setzt auf eine gemischte Zusammensetzung von internen und externen ExpertInnen. Dabei werden nur kleinere Projekte ausschließlich von internen ExpertInnen evaluiert. Auch beim RCN kommen sowohl interne als auch externe ExpertInnen zum Einsatz, wobei die internen GutachterInnen stärker auf formale und ökonomische Kriterien fokussieren. Ähnlich geht der deutsche Projektträger VDI/VDE-IT vor. Auch hier überprüfen die internen ExpertInnen zunächst einmal die formalen Kriterien, bevor es dann in der Hälfte der Fälle (abhängig von der Anzahl der Proposals) zum Einsatz von externen Gutachtern für den inhaltlichen Teil kommt. Der PT TÜV setzt ebenfalls auf ein gemischtes Vorgehen, wobei es vom Ministerium abhängt, ob intern oder extern evaluiert wird.

GutachterInnen	Fälle
Intern	Tekes PT VDI-TZ Vinnova (bis 2012)
Extern	FFG (Förderungswürdigkeit intern begutachtet)
Gemischt	IFD RCN Vinnova (seit 2013 zunehmend extern) PT VDI/VDE-IT PT TÜV PTJ

Tabelle 2: Einsatz interner und externer GutachterInnen

Interessant erscheint dabei, dass keine Forschungsförderungsagentur bei den externen Gutachten ausschließlich auf internationale ExpertInnen zurückgreift. Die Variationsbreite ist dabei aber durchaus beträchtlich: die FFG setzt meist internationale Gutachter ein, während die TACR nur im Rahmen von Kompetenzzentren auf diese zurückgreift und die RCN Innovation Division meist nationale GutachterInnen einsetzt.

Um zu verhindern, dass verschiedene GutachterInnen Beurteilungskriterien unterschiedlich interpretieren, kommen verschiedene Verfahren zum Einsatz. Besonders oft wurde der Versuch Kriterien möglichst klar zu formulieren und den GutachterInnen zu kommunizieren angeführt. Häufig wurden auch die Vorzüge von Panelsystemen hervorgehoben, da hier die Arbeit der GutachterInnen durch den Einsatz eines mehrfachen Begutachtungsverfahrens durch gegenseitige Überprüfung und Diskussion im ExpertInnengremium einer andauernden Angleichung ausgesetzt ist. Es wurde aber zum Beispiel von dem Interviewpartner beim PTJ angemerkt, dass es wichtig sei, auf die Stimmungslage in einem solchen Panel zu achten. So kann es vorkommen, dass die ExpertInnen am Anfang einer Sitzung großzügiger sind als am Abend (vergleiche Donovan 2007, Mallard et al 2009). Daher bietet es sich auch an, eigene Prozessverantwortliche in Begutachtungsverfahren mit Aufsichtsaufgaben zu betrauen. Es kommt aber auch vor, dass keine Meetings durchgeführt werden und die Kalibrierung dann von internen ExpertInnen über einen Vergleich der Gutachten durchgeführt wird (PT-TÜV).

Der vergleichenden Bewertung von Forschungsanträgen wird, unabhängig davon ob sie in Panels oder in anderer Form durchgeführt wird, eine hohe Bedeutung zugemessen, da man sich dadurch bessere Ergebnisse erhofft. Häufig bewertet bei dieser Vorgehensweise ein/e GutachterIn mehrere Anträge und ein Antrag wird von mehreren GutachterInnen evaluiert. Bei einer vergleichenden Vorgehensweise kommt es auch vor, dass die GutachterInnen die Bewertungen der anderen GutachterInnen lesen, bevor sie den Antrag in einer gemeinsamen Sitzung diskutieren. Beim PT VDI-TZ werden zum Beispiel alle Anträge von allen damit betrauten ExpertInnen begutachtet. Das Auswahlverfahren von TACR wiederum baut in besonderem Maße auf die vergleichende Bewertung von Anträgen: so finden sich in den meisten Programmen zwei aufeinanderfolgende Panels. Die einzige wirklich große Ausnahme bildet Tekes, wo die Proposals von Firmen keiner vergleichenden Begutachtung unterliegen. Ansonsten ist es üblich, nur Projekte mit relativ kleiner Fördersumme nicht im Vergleich zu beurteilen (IFD). Einen komplett anderen Weg hat der RCN eingeschlagen, wo versucht wird, mithilfe von Algorithmen und der Abstimmung einzelner Indikatoren für eine Kalibrierung der Expertise zu sorgen.

In den Forschungsförderungsorganisationen gibt es unterschiedlich intensive Bemühungen um eine Kalibrierung der Begutachtungen, die in Tabelle 3 wiedergegeben werden.

Kalibrierung	Fälle
Intensiv	RCN Vinnova FFG (abhängig von Programm) TACR (Panel System)
Mittel	Tekes PTJ IFD
Gering	PT VDI-TZ PT VDI/VDE-IT PT TÜV

Tabelle 3: Bemühungen um Kalibrierung von Begutachtungen

Begutachtung und Finanzierungsentscheidung sind bei der Forschungsfinanzierung üblicherweise getrennt. Meist liegt das Begutachtungsverfahren in den Händen interner und externer ExpertInnen, im IFD wird bei besonders großen Förderprogrammen auf Programmbeiräte zurückgegriffen. Während die formalen Finanzierungsentscheidungen häufig bei Ministerien liegen, werden die tatsächlichen Entscheidungen entweder vom Management der Forschungsförderungsagenturen (RCN, Tekes, Vinnova, TACR, teilweise IFD), von Beiräten, Panels und Jurys (FFG, teilweise IFD) und nur im Fall von Deutschland auf der Ebene der Referatsleitungen von Bundesministerien getroffen.

Fälle	Begutachtung	Finanzierungsentscheid
AT	Interne/externe Gutachter	Panels und Jury (Programm COIN, Thematische Programme), Beirat (Programm BRIDGE und Basisprogramme), selten Ministerien
DE	Projekträger: interne und externe Gutachter, (PTJ Externe, VDI/VDE-IT Interne)	Referatsleitung Ministerien
DK	IFD Interne/externe Gutachter	Management/Beirat der Förderagentur
CZ	Externe Gutachter, Beiräte	Management der Förderagentur
FI	Interne Gutachter	4 Managementebenen der Förderagentur
NO	Interne/externe Gutachter	3 Ebenen der Beirätestruktur der Förderagentur
SE	Interne/externe Gutachter	Management der Förderagentur

Tabelle 4: Begutachtung und Finanzierungsentscheid

Maßnahmen, die Lernen durch Erfahrung befördern sollen, unterscheiden sich stark voneinander. Hierbei fallen besonders die Unterschiede zwischen Deutschland und Tschechien beziehungsweise den übrigen untersuchten Agenturen auf. In den deutschen Projekträgern VDI-TZ und Jülich wird eher auf einer persönlichen Ebene aus Erfahrung gelernt. Neuen MitarbeiterInnen wird über ein Einarbeitungskonzept das nötige Wissen weitergegeben. Beim PT VDI-TZ wird verstärkt versucht, Wissen auch auf formalem Weg zu sammeln. Beim PT VDI/VDE-IT werden Erfahrungen in informellen Gesprächen ausgetauscht. Der PT-TÜV dokumentiert die Ergebnisse der Gutachten und der Diskussionsrunden. Die deutschen Projekträger setzen also auf persönliche Weitergabe von Wissen. Ähnlich ist es in der TACR, wo Wissen vorwiegend innerhalb der mit ExpertInnen besetzten Beiräte angesammelt wird. Diese sitzen in der Regel acht Jahre in dem Beirat und sammeln in dieser Zeit Wissen und Erfahrungen an, die aber nur selten formal festgehalten werden.

Ein besonderer Fall in Bezug auf Lernen aus Erfahrung bildet der RCN in Norwegen. Hier wurde das Projektauswahlverfahren 10 Jahre lang von WissenschaftlerInnen begleitet. Diese betrachteten die verwendeten Kriterien und den Output der vergebenen Förderungen und erarbeiteten so ein Modell auf welche Weise mit welchen Kriterien die Vergabe von Fördermitteln vorstangehen soll. Auch Vinnova bemüht sich auf formalem Weg aus Erfahrungen zu lernen, evaluiert dazu regelmäßig die internen Prozesse und untersucht die langfristigen Auswirkungen der Förderaktivitäten. Ähnlich die Vorgangsweise in Dänemark:

hier existiert die Praxis jedes Programm zweimal nach seinem Ende, im Abstand mehrerer Jahre, zu evaluieren. Zusätzlich dazu wurden bislang die langfristigen Ergebnisse der Programme mithilfe einer Datenbank analysiert, in der die Daten sämtlicher dänischer Unternehmen beinhaltet sind. Hierbei wurden geförderte mit nicht geförderten Firmen verglichen. Tekes lässt seine Programme regelmäßig evaluieren und es liegt eine Best-Practice Sammlung in Bezug auf die Auswahlverfahren und die anzuwendenden Kriterien der internen ExpertInnen vor, zudem gibt es Workshops für die GutachterInnen. Die FFG verfügt über ein internes Programmmonitoring, das einen Beitrag zum Lernen leisten soll und lässt ihre Programme regelmäßig durch externe GutachterInnen evaluieren.

Zusammengefasst zeigt sich also, dass verschiedene Lerninstrumente, wie beispielsweise ex-ante und ex-post Evaluationen, laufende Projektbegleitungen, TutorInnenprogramme für neue MitarbeiterInnen zur Anwendung kommen. Zusätzlich werden begleitende Untersuchungen von organisationalen Prozessen, wie etwa Projektauswahlverfahren, von den Agenturen in unterschiedlicher Weise eingesetzt, um die Wirkungen der vergebenen Förderungen zu verbessern. Die Veränderungen von organisationsbezogenen Prozeduren vieler Agenturen in den letzten Jahren im Zusammenhang mit dem Einsatz dieser Instrumente weist auf deren Einflusskraft hin (vgl. Biegelbauer 2013, 2016).

Tabelle 5 spiegelt die Bemühungen der Forschungsförderungsorganisationen, aus Erfahrung zu lernen, wider.

Intensität der Bemühungen	Fälle
Intensiv	RCN Vinnova IFD FFG
Mittel	Tekes PTJ PT VDI-TZ
Gering	PT VDI/VDE-IT PT TÜV TACR

Tabelle 5: Bemühungen aus Erfahrung zu lernen

ABSCHLIESSENDE BETRACHTUNGEN

Der vorliegende Beitrag beschäftigte sich mit Projektauswahlverfahren in Forschungsförderungsagenturen. In einem Vergleich wurden der Innovation Fund Denmark (IFD), der Research Council of Norway (RCN) in Norwegen, die Verket för innovationssystem (Vinnova) in Schweden, Teknologian kehittämisskeskus (Tekes) in Finnland, die österreichische Forschungsförderungsgesellschaft (FFG), die Technologicke agentury BR (TACR), sowie die deutschen Projektträger Jülich (PTJ), Technischer Überwachungsverein Rheinland (PT TÜV Rheinland), Verein Deutscher Ingenieure/Technologiezentrum (PT VDI-TZ) und der Verein Deutscher Ingenieure/Verein Deutscher Elektrotechniker-Innovation und Technik (PT VDI/VDE-IT) untersucht. Abschließend werden zuerst einige organisatorische Praktiken herausgestellt, die als modelhaft gelten können. Daran

anschließend werden einige Trends präsentiert, welche die Entwicklungsrichtungen der einzelnen Organisationen beschreiben.

Einige **modellhafte organisationale Praktiken** sind in den untersuchten Fällen zu Tage getreten und wurden von verschiedenen GesprächspartnerInnen auch immer wieder als bedeutsam bezeichnet.

- Die Klarheit der Bedeutung der verwendeten Kriterien für die Projektauswahl. Die Kriterien sollten nicht nur schon bei der Entwicklung der Programme festgelegt, sondern auch gegenüber AntragstellerInnen, GutachterInnen und AgenturmitarbeiterInnen kommuniziert werden. Die Benotung der einzelnen Kriterien sollte sich nicht in einem zahlengestützten Schema erschöpfen, sondern die einzelnen Noten sollten jeweils einem Satz zugeordnet sein, der eine bestimmte Bewertung zum Ausdruck bringt.
- Der Einsatz externer wie interner GutachterInnen. Die meisten Organisationen setzen externe und interne GutachterInnen für verschiedene Funktionen im Rahmen der Projektauswahlverfahren ein. Die spezifischen Funktionen für die beiden Gruppen unterscheiden sich im Einzelfall. Vor allem bei größeren Ausschreibungen mit höherer wissenschaftlicher Spezialisierung werden häufig externe GutachterInnen eingesetzt, die durch organisationsinterne ExpertInnen ergänzt werden, die Formalprüfungen und oft auch betriebswirtschaftliche Prüfungen durchführen.
- Konzentration auf die „mittelguten“ Anträge. In vielen Organisationen werden Entscheidungen über Projektanträge in zwei Stufen durchgeführt: die besonders guten und die besonders schlechten Anträge werden aus der Erfahrung, dass über diese beiden Kategorien rasche Übereinstimmung hergestellt werden kann, in einem ersten Schritt aussortiert. In einem zweiten Schritt werden die mittleren Anträge analysiert und diskutiert, wodurch mehr Zeit für diese, im Hinblick auf ihr Auswahlkriterien-bezogenes Abschneiden, weniger klaren Fälle bleibt.
- Kalibrierung von Bewertungen. In den meisten Organisationen gibt es ein Augenmerk auf die potentiell unterschiedliche Auslegung einzelner Kriterien durch die GutachterInnen, die einen spezifischen Hintergrund mitbringen, der sie Kriterien in unterschiedlicher Art und Weise interpretieren lässt. Eine häufig angewendete Methode, um mit dem Problem ungleicher Gewichtung umzugehen, ist die vergleichende Beurteilung von Projektanträgen, innerhalb derer einerseits jede/r GutachterIn mehrere Anträge zugeteilt bekommt und andererseits jeder Antrag durch mehrere GutachterInnen begutachtet wird. Dazu verwenden die meisten Forschungsförderungsorganisationen Panels, in denen die einzelnen GutachterInnen ihre Bewertungen mit KollegInnen offenlegen und diskutieren müssen. Programmkomitees sind eine verstetigte Form dieser Panels.
- Lernen aus Erfahrung. In vielen Organisationen gibt es Mechanismen, um Erfahrungen zu sammeln, einzuschätzen und für Veränderungen zu verwenden. Dazu werden beispielsweise Debriefing-Sitzungen nach dem Ende von Auswahlverfahren verwendet, die zu Good Practice Beschreibungen führen, die

wiederum in Personalentwicklungsaktivitäten Eingang finden können oder von Prozessentwicklungen aufgegriffen werden. Programmkomitees können hier eine wesentliche Rolle spielen. Oftmals extern gebildetes Reflexionswissen aus Evaluationen und begleitenden Untersuchungen von organisationalen Prozessen ergänzen das Erfahrungswissen.

In einer Zusammenschau der Ergebnisse lassen sich in verschiedener Hinsicht die neun untersuchten Forschungsförderungsorganisationen in drei Gruppen einteilen, von denen eine die Förderorganisationen in Dänemark, Finnland und Norwegen, eine zweite die Agenturen in Schweden und Österreich, sowie eine dritte, die Projektträger in Deutschland umfasst. So sind die deutschen Projektträger älter als die anderen untersuchten Organisationen, sie sind darüber hinaus diverser in Bezug auf ihre Budgetgrößen, die Projektauswahlverfahren, den Einsatz interner bzw. externer GutachterInnen sowie der Prozeduren, die sich auf das Lernen aus Erfahrung beziehen. Die letzte Eigenschaft und die im Vergleich zu den anderen Organisationen größere Abhängigkeit von den Ministerien sind auch Elemente einer Erklärung der geringeren Bereitschaft Veränderungen an internen Prozeduren durchzuführen.

Über diese Differenzierung hinaus lassen sich jedoch an den untersuchten Forschungsförderungsorganisationen einige **Institutionen-übergreifende Entwicklungstrends** feststellen:

- Angleichung der Projektauswahlverfahren. Während es zwischen den einzelnen Forschungsförderungsorganisationen Unterschiede in Bezug auf die Projektauswahlverfahren gibt, gibt es innerhalb der einzelnen Organisationen einen Trend zur Angleichung der jeweiligen Verfahren. Dieser ist in den nordischen Ländern deutlich ausgeprägt, wo Tekes ein Standardverfahren verwendet, RCN 2011 ein derartiges eingeführt hat, der IFD weitgehend vereinheitlichte Verfahren anwendet und Vinnova ausgehend von einem als besonders erfolgreich angesehenen Programm dabei ist Elemente in den Auswahlverfahren anzugleichen. In deutschen Projektträgern gibt es im Gegensatz dazu eine Vielfalt unterschiedlicher Projektauswahlverfahren, wobei einzelne Bereiche, etwa im PT VDI-TZ, in Zusammenarbeit mit den jeweiligen ministeriellen Referaten eine gewisse Vereinheitlichung erzielt haben.
- Einsatz externer GutachterInnen. Ein weiterer Trend ist der Einsatz externer GutachterInnen, auf die in mehreren Organisationen in größerem Ausmaß zurückgegriffen wird, als das noch vor einigen Jahren der Fall war. So werden etwa in der Vinnova externe GutachterInnen stärker eingesetzt, mit dem Ziel die Begutachtungsprozesse in den nächsten Jahren größtenteils auf externe GutachterInnen umzustellen. Auch im IFD und in Teilen des deutschen PT Jülich werden deutlich mehr externe GutachterInnen eingesetzt, als das noch vor kurzem der Fall war.
- Objektivierung. In einer Reihe von Organisationen werden Anstrengungen unternommen die Projektauswahlverfahren zu objektivieren. Im RCN kommt seit zwei Jahren in der gesamten Organisation ein mehrstufiger Kriterien gestützter Prozess der Projektauswahl zur Anwendung, die in der Vinnova im Moment laufende Prozessentwicklung hat eine Objektivierung durch

eine neue Kombination unterschiedlicher GutachterInnen sowie den Einsatz eines Programmkomitees zum Ziel. In der TACR wird in diesem Zusammenhang besonders stark auf dem Einsatz von Beiräten gesetzt. Diese wurde in der Vergangenheit allerdings reformiert. So können die beiden Gremien jeweils nur mehr 10 Prozent der vergebenen Punkte pro Antrag ändern, um sicher zu stellen, dass die Beiräte tatsächlich nur eine objektivierende Funktion einnehmen.

- Erfahrungslernen. Beinahe alle Organisationen haben im Verlauf der letzten Jahre die Anzahl der Instrumente, die ein Sammeln und Bewerten von Erfahrungswissen erlauben, vergrößert. Diese sind bei der Vinnova und im RCN besonders vielfältig und umfassen dort etwa den Versuch einer laufenden Prozessentwicklung mit regelmäßigen prozeduralen Veränderungen im Verlauf der letzten Jahre auf der Basis der Sammlung von Erfahrungswissen, Systemevaluationen, eine durchgehende Begleitung der Untersuchung von organisationalen Prozeduren und ein System von standardisierten Evaluationen zu verschiedenen Punkten im Verlauf eines Programmes.

BIBLIOGRAPHIE

Arrow, K. J. (1971, 1.ed 1962): Economic Welfare and the Allocation of Resources for Invention. In: The Economics of Technological Change - Selected Readings. N. Rosenberg. Harmondsworth: Penguin Books, S. 164-181.

Biegelbauer, P. (2000): 130 Years of Catching Up With the West: A Comparative Perspective on Hungarian Industry, Science and Technology Policy-Making Since Industrialization. Aldershot: Ashgate Publishing.

Biegelbauer, P. (2013): Wie lernt die Politik - Lernen aus Erfahrung in Politik und Verwaltung. Wiesbaden: VS Verlag für Sozialwissenschaften.

Biegelbauer, P. (2016, forthcoming): "How Different Forms Of Policy Learning Influence Each Other: Case Studies from Austrian Innovation Policy-Making." Policy Studies 37.

Biegelbauer, P. und T. Palfinger (2014): "Verschiedene Verfahren der Auswahl von Forschungsprojekten: ein Vergleich von neun angewandten Forschungsförderungsorganisationen." Forschung 7(1+2): 46-51.

Braun, D. (1997): Die politische Steuerung der Wissenschaft. Frankfurt am Main: Campus.

Christensen, T. und P. Laegreid, Eds. (2006): Autonomy and Regulation: coping with agencies in the modern state. Cheltenham: Edward Elgar.

Commission, European (2013): Innovation Union Scoreboard 2013: Enterprise and Industry. Inno Metrics. Brüssel: European Commission.

Commission, European (2015): Innovation Union Scoreboard 2015: Brüssel: European Commission.

Dinges, M. (2010): „Öffentliche FTI-Projektfinanzierung in Österreich:

Ausmaß und Bedeutung im Innovationssystem.“ In: Steuerung von Wissenschaft? Die Governance des österreichischen Innovationssystems. P. Biegelbauer. Innsbruck: Studienverlag, S. 109-141.

Donovan, C. (2007): „Introduction: Future pathways for science policy and research assessment: metrics vs peer review, quality vs impact.“ *Science and Public Policy* 34(8): 538-542.

Kubeczko, K. und K. M. Weber (2009): „Proaktive Legitimation FTI-politischer Interventionen: Jenseits von Markt- und Systemversagen.“ In: *Innovationsforschung und Technologiepolitik in Österreich*. K.-H. Leitner, K. M. Weber und J. Fröhlich. Innsbruck: Studienverlag, S. 127-153.

Lepori, B., P. van der Besselaar, M. Dinges, B. van der Meulen, B. Potí, E. Reale, S. Slipersaeter und J. Theves (2007): "Indicators for Comparative Analysis of Public Project Funding. Concepts, implementation and evaluation." *Research Evaluation* 16(4): 243-255.

Mallard, G., M. Lamont, J. Guetzkow (2009): "Fairness as Appropriateness: Negotiating Epistemological Differences in Peer Review." *Science, Technology & Human Values* 34(5): 573-606.

Weber, K. M. (2009): „FTI Politik im Spiegel von Theorie und Praxis: von Planung über Steuerung zu Governance.“ In: *Innovationsforschung und Technologiepolitik in Österreich*. K.-H. Leitner, K. M. Weber und J. Fröhlich. Innsbruck, Studienverlag, S. 231-254.

INTERNET

<http://www.forskningsradet.no/en/Organisation/1138785841802>, abgerufen am 3.11.2015.

<http://innovationsfonden.dk/en/about-ifd>, abgerufen am 3.11.2015 http://www.forskningsradet.no/en/Key_figures/1138785841814, abgerufen am 3.11.2015.

http://www.photonikforschung.de/fileadmin/MEDIENDATENBANK/Dokumente/Broschueren/BMBF_Photonik_Forschung_Deutschland_final_1.pdf, abgerufen am 3.11.2015.

http://www.photonikforschung.de/fileadmin/MEDIENDATENBANK/SERVICE/Publikationen/Agenda_Photonik2020_11-2010.pdf, abgerufen am 3.11.2015.

<http://www.ptj.de/auftraggeber>, abgerufen am 3.11.2015.

<http://www.ptj.de/geschaeftsbericht>, abgerufen am 3.11.2015.

<http://www.ptj.de/geschaeftsfelder>, abgerufen am 3.11.2015.

<http://www.ptj.de/profil>, abgerufen am 3.11.2015.

<https://www.tacrcz/index.php/en/about-tacrc.html>, abgerufen am 3.11.2015

<http://www.taftie.org/content/tekes-finnish-funding-agency-innovation-finland>, abgerufen am 3.11.2015.

<http://www.taftie.org/content/vdivde-it-germany>, abgerufen am 3.11.2015.

http://www.tuvpt.de/fileadmin/pdf/Dokumentationen/verkehrsfor-schungsprogramm-mobilitaet-und-verkehrstechnologien_property_pdf_bereich_bmw_i_sprache_de_rwb_true.pdf, abgerufen am 3.11.2015.

http://www.vditz.de/fileadmin/media/news/images/VDI_Technologiezentrum.pdf, abgerufen am 3.11.2015

<http://www.vditz.de/service/unser-profil/>, abgerufen am 3.11.2015

<http://www.vditz.de/forschungsfoerderung/>, abgerufen am 3.11.2015.

<http://www.vdivde-it.de/wir-ueber-uns>, abgerufen am 3.11.2015.

<http://www.vdivde-it.de/wir-ueber-uns/team>, abgerufen am 3.11.2015.

<http://www.vinnova.se/en/About-VINNOVA/>, abgerufen am 3.11.2015.

<http://www.vinnova.se/en/About-VINNOVA/Organization-and-management/>, abgerufen am 3.11.2015.

<http://www.vinnova.se/en/Our-activities/>, abgerufen am 3.11.2015.

<http://www.vinnova.se/en/Our-activities/Cross-border-co-operation/Challenge-driven-Innovation/>, abgerufen am 3.11.2015.

AUTOREN

PETER BIEGELBAUER

AIT Austrian Institute of Technology, Innovation Systems Department

E: peter.biegelbauer@ait.ac.at

THOMAS PALFINGER

AIT Austrian Institute of Technology, Innovation Systems Department

E: tpalfinger@gmx.at

„EVALUATION VS. EDUCATED GUESS“ PROJEKTTRÄGER ALS UNTERSCHÄTZE QUELLE FÜR DIE BEWERTUNG VON FÖRDERPROGRAMMEN

JAN WESSELS, CHRISTIANE KERLEN UND SONJA KIND

EINLEITUNG

Evaluationen setzen sich als verpflichtende Standardelemente einer Überprüfung und Bewertung von Fördermaßnahmen zunehmend durch (vgl. Böttcher et al. 2014). Sie ergänzen bzw. ersetzen dabei die Bewertung durch die Programmverantwortlichen (Ministerien und Projektträger), die aufgrund ihrer Fachexpertise und dem direkten Kontakt mit den Geförderten zuvor allein die Bewertung ihrer Programme vornahmen (Astor et al. 2014).

Als Begründung für diesen Wandel werden sowohl eine 1) unabhängigere als auch eine 2) bessere (also zutreffendere, genauere) Bewertung durch externe Evaluatorinnen und Evaluatoren angeführt. Evaluierende sollen als unabhängige Dritte die Programmumsetzung auf ihre Zielerreichung hin objektiver prüfen, da angenommen werden kann, dass sie kein Eigeninteresse an bestimmten Ergebnissen haben. Aufgrund ihrer Methodenexpertise sollen sie zudem hochwertigere Steuerungsinformationen für die Gestaltung von Fördermaßnahmen liefern können.

Von dieser Rolle der unabhängigen Bewerter bewusst unterschieden werden die Projektträger. Die Bundesregierung beschrieb im Rahmen einer Antwort auf eine kleine Anfrage über die enge Involvierung der Projektträger in die Programmevaluationen knapp (Deutsche Bundesregierung 2011):

„Die Projektträger sind bei der Evaluierung von Förderprogrammen nur unterstützend tätig und evaluieren die Programme nicht selbst. Diese Praxis soll auch in Zukunft beibehalten werden.“

Es stellt sich die Frage, wie sich das Wissen der Programmverantwortlichen von den gewonnenen empirischen Daten durch externe Evaluierende unterscheidet. Der direkte Vergleich zwischen den Ergebnissen einer extern beauftragten Evaluation und des auf vielfältigen Quellen und Einschätzungen beruhenden „educated guess“ der Programmverantwortlichen zu spezifischen Erhebungsdimensionen soll daher in diesem Beitrag beleuchten, an welchen Stellen Evaluationsergebnisse zusätzlichen Erkenntnisgewinn für die Programmverantwortlichen /

Projektträger bieten, und an welchen Stellen ihre eigene Kenntnis von Programm und Programmakteuren genutzt werden könnte.

Externe Evaluationsteams verfügen in der Regel über besondere Feldkenntnisse, das heißt sie kennen das (Politik-/Technologie-)Feld, in dem das Programm durchgeführt wird, durch ihre Evaluationstätigkeit in diesem Bereich. Sie sind jedoch nur in Ausnahmefällen selbst Fachexperten im engeren Sinne, also inhaltliche Experten für das spezifische Fachprogramm, das evaluiert werden soll.¹ Nicht selten wird der Mehrwert der externen Evaluation daher von den Programmverantwortlichen angezweifelt. Als Argument wird dann neben dem fehlenden Expertenstatus angeführt, dass die Programmverantwortlichen selbst fundierteres Wissen über die Programmakteure haben, als in einer externen Evaluation erlangt werden könne. Hier handelt es sich in der Regel um im Begleitprozess der Projekte erworbenes Wissen, bei dem nicht mehr rekonstruiert werden kann, auf welche Weise Schlussfolgerungen oder Bewertungen abgeleitet wurden (ein so genanntes „Bauchgefühl“; vgl. Gigerenzer/Gaissmaier 2015). In diese Kategorie fällt auch Erfahrungswissen, welches über eine längere Tätigkeit als Projektträger über mehrere Projektträgerschaften hinweg erworben wurde und für generalisierbare Charakteristika eines Förderprogramm-Typs (z.B. eines Technologieförderprogramms) gilt.

Projektträger akzeptieren Evaluationen daher häufig nur aufgrund einer extern auferlegten Verpflichtung. Dazu trägt bei, dass ihr Wissen im Rahmen von Evaluationen selten systematisch genutzt wird. In der Regel geht ihre Einbindung über die Zulieferung von Projektdaten und so genannten Prozessdaten (Dauer des Antragsverfahrens, Zahlungsströme, etc.) sowie Interviews zur internen Implementierungspraxis innerhalb der Projektträger nicht hinaus. Ihre Bewertung von Projekten und Programm wird kaum systematisch erhoben. Dies führt dazu, dass sie an der Bewertung „ihres Programms“ nur selten beteiligt sind, auch weil eine Befangenheit vermutet wird. Letztlich kann diese Distanz aber dazu führen, dass die Evaluationsergebnisse von den Projektträgern nicht akzeptiert und daher nicht umfassend genutzt werden. Hierfür gibt es zurzeit allerdings nur anekdotische Evidenz aus Hintergrundgesprächen mit Projektträgerverantwortlichen und Fachreferaten, die auf eine nicht

1 Zum Teil werden Evaluationsteams durch externe Fachexperten im Unterauftrag ergänzt. Zuweilen steht der Zusammenarbeit jedoch eine Befangenheit zum Thema entgegen.

unerhebliche Reserviertheit schließen lassen. Im Sinne einer besseren Evaluationskultur wäre es aus Sicht der Autoren wünschenswert, diese Distanz ein wenig abzubauen.

Der vorliegende Beitrag wirft Licht auf diese Problemstellung, indem er auf der Basis eines empirischen Experiments die Ergebnisse von Evaluationserhebungen und die Einschätzungen von Programmverantwortlichen miteinander vergleicht. Es wird diskutiert, welche Themen und Fragestellungen sich besonders für eine Bewertung durch die Programmverantwortlichen eignen und wie dies dazu führen kann, möglichst effiziente Verfahren der Datengenerierung einzusetzen, ohne auf hochwertige Evaluationsergebnisse zu verzichten. Nicht zuletzt wird auch die Frage aufgeworfen, welche Konsequenzen sich daraus für eine Kultur der Evaluation ergeben.

VORGEHEN

Das empirische Experiment wurde im Rahmen einer Programmevaluation eines Technologieförderprogramms durchgeführt. Es handelt sich um die Evaluation eines Programms im Bereich Elektromobilität. Das verantwortliche Ministerium zielt mit diesem Programm insbesondere auf die Förderung neuer Technologien und -verfahren, die zu einer schnellen Überführung von Innovationen in vermarktbare Komponenten, Fahrzeuge und Systeme beitragen können. Aufgrund der guten Zusammenarbeit mit dem auftraggebenden Referat im Ministerium und dem Projektträger war es möglich, dieses Experiment in Ergänzung zum ursprünglichen Evaluationsdesign durchzuführen.

Als Programmverantwortliche wurden Mitarbeiter des Projektträgers (die Einzelprojektbetreuer sowie der für das gesamte Programm Verantwortliche) gebeten, das Antwortverhalten der Projektnehmer vorab einzuschätzen und auch anzugeben, wie sicher bzw. unsicher sie sich in dieser Einschätzung fühlen. Diese Einschätzungen wurden dann mit dem tatsächlichen Antwortverhalten der Projektnehmer verglichen.

Für das Experiment wurden drei Themenkomplexe ausgewählt, bei denen davon ausgegangen werden kann, dass die Programmverantwortlichen einzelne Projekte, beispielsweise durch die Begleitung während der Antragstellung und das Lesen der Anträge, gut kennen:

- die Ziele der einzelnen Vorhaben,
- die Abschätzung von Input-Additionalität/ Mitnahmeeffekten der Förderung sowie
- die Beurteilung des Antragsverfahrens.

Für das empirische Experiment wurden in diesem Fall nur die Fragen aus dem Online-Fragebogen an die Projektnehmer herangezogen. Weitere Datenquellen zur Analyse und Bewertung der hier betrachteten Themenkomplexe, die im Rahmen des Multimethodendesigns der Programmevaluation genutzt wurden (z.B. Angaben zur Dauer der Bearbeitung aus den Förderakten, Vergleichszahlen aus anderen Programmevaluationen, Auskünfte zu den genannten Fragen von Projektträgermitarbeitenden und Projektnehmern in Interviews) wurden bei dem Experiment nicht weiter berücksichtigt. Die ausgewählten Fragen aus dem Online-Fragebogen wurden den Mitarbeitern des Projektträgers übermittelt, mit der Bitte

- a. das Antwortverhalten der Projektnehmer zu antizipieren („Bitte kreuzen Sie für jede Einzelfrage an, welches Antwortverhalten Sie erwarten“)
- b. die eigene Sicherheit in Bezug auf dieses antizipierte Antwortverhalten anzugeben („Kreuzen Sie bei Bedarf an: - bin mir sehr unsicher / - kann ich nicht einschätzen“)

Alle Projektbetreuer wurden gebeten, je einen Fragebogen zu den von ihnen betreuten Vorhaben auszufüllen. Außerdem wurde der für das gesamte Programm verantwortliche Mitarbeiter des Projektträgers gebeten, einen Fragebogen auszufüllen, welche Antworten er im Durchschnitt von allen Projekten erwartet.

Die Analyse erfolgte im Hinblick auf die Treffgenauigkeit der Antworten, die Streuung des Antwortverhaltens und die selbsteingeschätzte Unsicherheit in Bezug auf die Simulation des Antwortverhaltens. Zur Untersuchung der Treffgenauigkeit wurde die Übereinstimmung zwischen Projektnehmerantworten – herangezogen wurden die Antworten der geförderten Unternehmen auf Ebene der Teilvorhaben (N=35) – und der Einschätzung des Antwortverhaltens durch den Projektträger ermittelt. Zur Analyse der Güte der Einschätzungen wurden die Antworten der Einzelbetreuer (N=11), die sich auf die Ebene der Verbundvorhaben beziehen, mit den Antworten des verantwortlichen Mitarbeiters für das gesamte Programm verglichen, um zu prüfen, ob es sinnvoll ist, die einzelnen Projektbetreuer zu befragen, oder ob es ausreichen würde, nur eine Person zu befragen. Zur Analyse von Unsicherheit wurde die Übereinstimmung mit der Unsicherheit der Einschätzung verglichen, um der Frage nachzugehen, wie gut die Projektträger ihre Kenntnisse selbst einschätzen können.

In der nachfolgenden Auswertung wurde jeweils ein Mittelwert für das Gesamtprogramm gebildet, also keine 1:1 Gegenüberstellung der Projektnehmerantworten und Projektträgerereinschätzungen auf Verbundprojektebene vorgenommen. Insgesamt ist die Streuung der Antworten der Projektnehmer größer als die Streuung der Antworten der Projektträgermitarbeiter. Dies wird im Folgenden außer Acht gelassen, da sich die Einschätzungen der Unternehmen auf ihr Teilvorhaben beziehen, während die Einschätzungen der Einzelbetreuer pro Verbundprojekt erfolgten.

FRAGESTELLUNGEN

Folgende Fragestellungen standen im Mittelpunkt der Untersuchung:

1. Durch den Vergleich der Antworten von Projektnehmern selbst mit der Einschätzung von Programmverantwortlichen zum Antwortverhalten der Projektnehmer soll zunächst untersucht werden, wie gut die Programmverantwortlichen eine Vorhersage über die Antworten von Projektnehmern machen können.
2. Durch den Vergleich der Antworten der Einzelbetreuer mit der Antwort des Gesamtverantwortlichen für alle Projekte soll geklärt werden, ob eine übergreifende Einschätzung des Projektträgers hinreichend ist oder ob die Antwortqualität steigt, wenn jeder fachliche Betreuer einzeln antwortet.

3. Außerdem soll untersucht werden, welche Themenkomplexe sich für die Bewertung durch Programmverantwortliche eignen. Für das empirische Experiment wurden drei Themenkomplexe (Projektziele, Antragsverfahren, Verhalten bei Nichtförderung) ausgewählt, die prinzipiell geeignet erscheinen, weil von einer guten Kenntnis der Projekte bei den Programmverantwortlichen ausgegangen werden kann.
4. Letztlich soll damit die Frage beantwortet werden, an welchen Stellen es in Programmevaluationen effizient und effektiv ist, die Programmverantwortlichen in das Evaluationsdesign mit einzubeziehen, weil sie genaue Angaben zu den einzelnen Projekten machen können. Auf diese Weise könnte z.B. der Erhebungsaufwand bei den Projektnehmern reduziert werden, ohne Qualitätsverluste für die Evaluation hinnehmen zu müssen, oder Erhebungen bei den Projektnehmern könnten auf einem alternativen Weg validiert werden.
5. Nicht zuletzt soll der Frage nachgegangen werden, ob die damit einhergehende Stärkung der Rolle der Programmverantwortlichen und die Würdigung ihres Wissensbestands eine Veränderung der Evaluationskultur nach sich ziehen könnte oder müsste.

tatsächlichen weitgehend überein. In der Regel wird die Bedeutung der Ziele von Seiten des Projektträgers nur etwas weniger hoch eingeschätzt als durch die Unternehmen. In einzelnen Aspekten sind jedoch auch Abweichungen zu verzeichnen. Die größte Abweichung besteht in Bezug auf die Bedeutung des Ziels „Erweiterung des Grundlagenwissens“. Dieser Aspekt erreicht in der Befragung der Unternehmen einen Wert rund „2“ auf einer Skala von 1 bis 6 (von 1 - sehr wichtiges Ziel bis 6 – unwichtiges Ziel), während die Projektbetreuer davon ausgingen, dass dieser Aspekt in anwendungsorientierten FuE-Vorhaben von etwas geringerer Bedeutung für Unternehmen ist (rund „3,5“). Dies trifft auch auf das Ziel „Bestehende Technologien verbessern“ zu, dass von den Unternehmen etwas wichtiger (rund „2“) als von den Projektverantwortlichen eingeschätzt wurde (rund „3“). Lediglich in zwei Dimensionen erwarteten die Projektverantwortlichen, dass diese Ziele für die Unternehmen deutlich wichtiger wären, als es tatsächlich der Fall war. Die Einschätzung zur Bedeutung der Ziele „Entwicklung neuer Technologien“ und insbesondere „Entwicklung eines Funktionsmusters“ wichen am deutlichsten ab.

Für diese Befragungsdimension konnte insgesamt eine relativ hohe Übereinstimmung von antizipiertem und tatsächlichem Antwortverhalten erwartet werden, da es bei der Zielstellung um bereits im Antrag beschriebene Sachverhalte ging. Gleichwohl war ebenso anzunehmen, dass in der Befragungssituation selbst der Projektantrag nicht direkt zum Ausfüllen des Fragebogens zur Hand genommen wurde, sondern diese Fragekategorien aus der Erinnerung heraus bzw. aus dem eigenen Projektverständnis heraus bearbeitet wurden. Die Abweichungen zwischen Unternehmensantworten und durch den Projektträger erwartetem Antwortverhalten spiegeln also möglicherweise leichte Interpretationsunterschiede in der Einschätzung der geförderten Projekte wieder, ohne das festgelegt werden könnte, welcher Akteur (gefördertes Unternehmen oder Projektträger) qua Rolle / Funktion eher „Recht“ hat.

ERGEBNISSE

TREFFGENAUIGKEIT

Im ersten Themenkomplex ging es um die Einschätzung dazu, welche Ziele die Unternehmen mit ihrer Beteiligung an den Teilvorhaben verfolgen. Wie Abbildung 1 zeigt, stimmen die antizipierten Antworten mit den

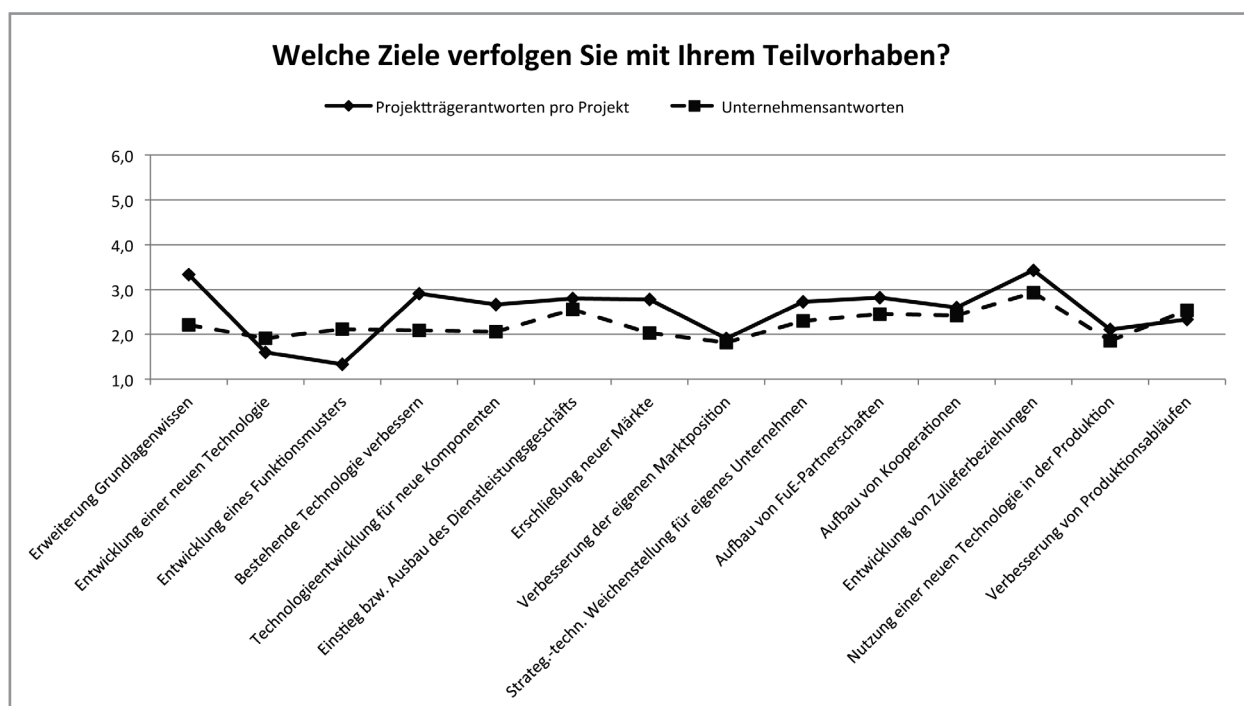


Abbildung 1: Angenommene Antworten des Projektträgers und tatsächliche Antworten der Unternehmen zu der Frage „Welche Ziele verfolgen Sie mit Ihrem Teilvorhaben?“

Noch größere Abweichungen waren bei der kontrafaktischen Frage („Wie hätten Sie bei einer Ablehnung des Projektes reagiert?“) zu verzeichnen (siehe Abbildung 2). Hier ging der Projektträger davon aus, dass die Unternehmen das Projekt auch ohne die Förderung durchgeführt hätten. Über 50% der Unternehmen gaben jedoch an, dass das Projekt ohne die Beteiligung an der Förderinitiative nicht durchgeführt worden wäre. Entsprechend abweichend sind die alternativ genannten Möglichkeiten, wenn das Projekt doch durchgeführt worden wäre, jedoch z.B. im Projektvolumen reduziert, mit verlängerter Laufzeit oder mit

veränderten Zielen. Hier gehen die Projektverantwortlichen eher davon aus, dass die Unternehmen das Projekt bei einem negativem Förderbescheid in eingeschränkter Form durchgeführt hätten. Bei der Frage, ob das Projekt ohne Veränderungen auch ohne Fördermittel durchgeführt worden wäre, gehen die Projektverantwortlichen zuversichtlich davon aus, dass dies nicht der Fall ist und es keine Mitnahmeeffekte gibt. Es zeigt sich jedoch, dass ein kleiner Teil der Unternehmen ihr Teilvorhaben nach eigener Einschätzung auch unabhängig von der Teilnahme am Förderprogramm umgesetzt hätte.

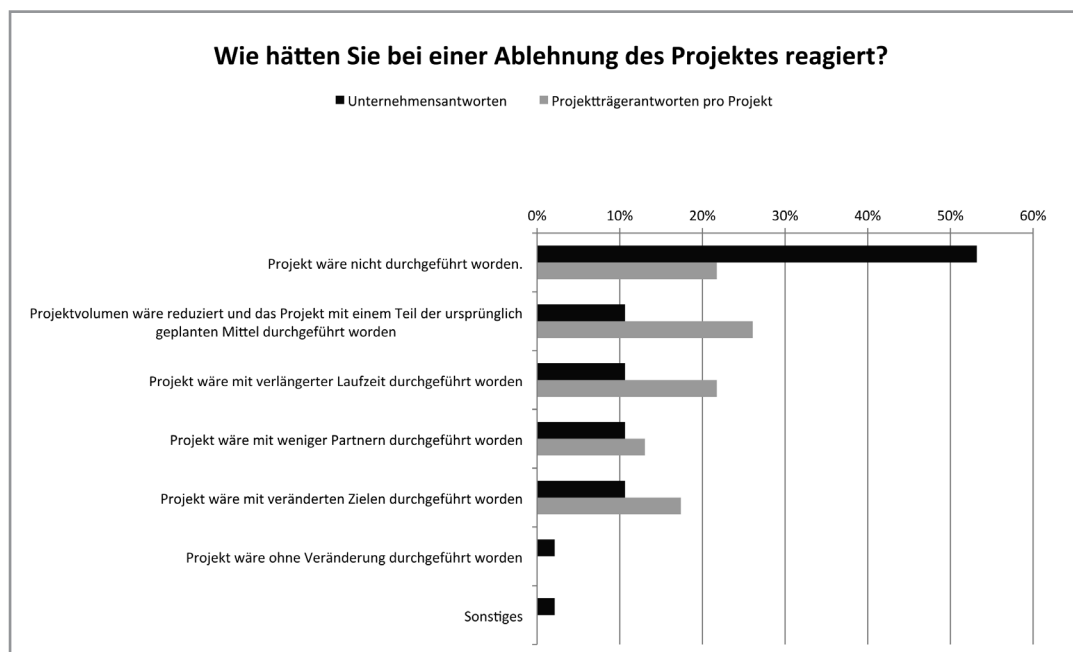


Abbildung 2: Angenommene Antworten des Projektträgers und tatsächliche Antworten der Unternehmen zu der Frage „Wie hätten Sie bei einer Ablehnung des Projektes reagiert?“

GÜTE DER EINSCHÄTZUNGEN JE NACH ROLLE BEIM PROJEKTTÄGER

Im nächsten Abschnitt wird der Frage nachgegangen, inwieweit die übergreifende Einschätzung des Gesamtverantwortlichen beim Projektträger von der individuellen Einschätzung der jeweiligen fachlichen

Betreuer abweicht. Können die fachlichen Betreuer der Projekte das Antwortverhalten der Unternehmen in „ihren“ Projekten möglicherweise besser als der Gesamtverantwortliche einschätzen?

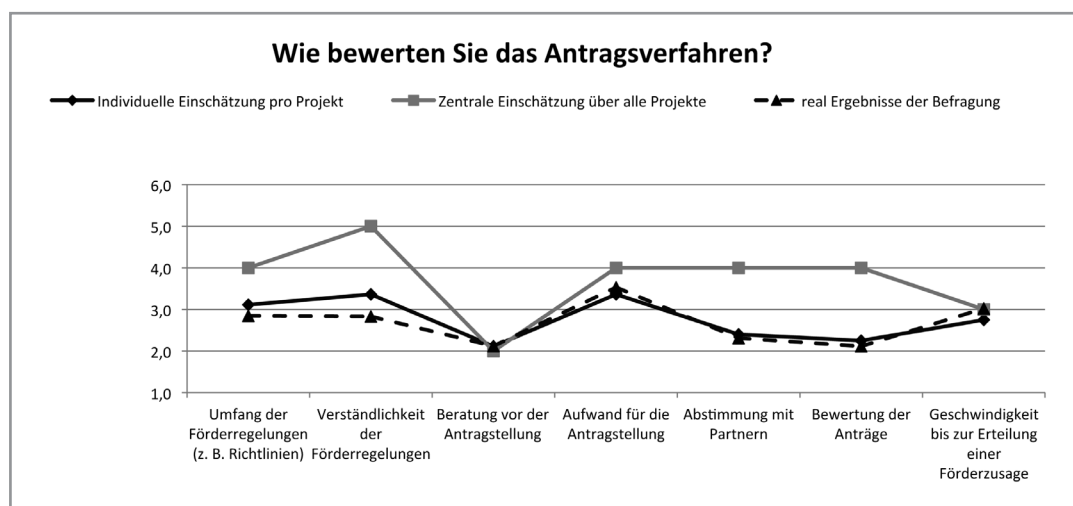


Abbildung 3: Angenommene Antworten des Projektträgers und tatsächliche Antworten der Unternehmen zu der Frage „Wie bewerten Sie das Antragsverfahren?“

Diese Frage wurde in Bezug auf die Bewertung des Antragsverfahrens untersucht (siehe Abbildung 3). Antworten erfolgten auf einer Skala von 1 bis 6 (von 1 - sehr angemessen bis 6 – gar nicht angemessen). Tatsächlich können die fachlichen Betreuer sehr genau antizipieren, wie die Unternehmen die einzelnen Aspekte und Schritte im Antragsverfahren bewerten. Im Gegensatz dazu weicht die Einschätzung des Gesamtverantwortlichen ab und fällt bis auf die beiden Dimensionen „Beratung vor der Antragstellung“ und „Geschwindigkeit bis zur Erteilung des Zuschlags“ durchgehend negativer aus. Dies zeigt, dass die fachlichen Betreuer sehr eng an „ihren“ Projekten sind und am besten einschätzen

können, wie die Unternehmen das gesamte Antragsverfahren bewerten.

Die Einschätzung durch den Gesamtverantwortlichen auf Projektträgerseite ist vergleichsweise pessimistisch und geht von einer negativeren Bewertung der Arbeit seiner Kolleginnen und Kollegen durch die Fördernehmer aus als tatsächlich der Fall ist. Vermutlich sind hier längerfristige Erfahrungswerte in die Einschätzung eingeflossen, die aus generalisierbaren Erkenntnissen aus der Betreuung von Förderprogrammen resultieren

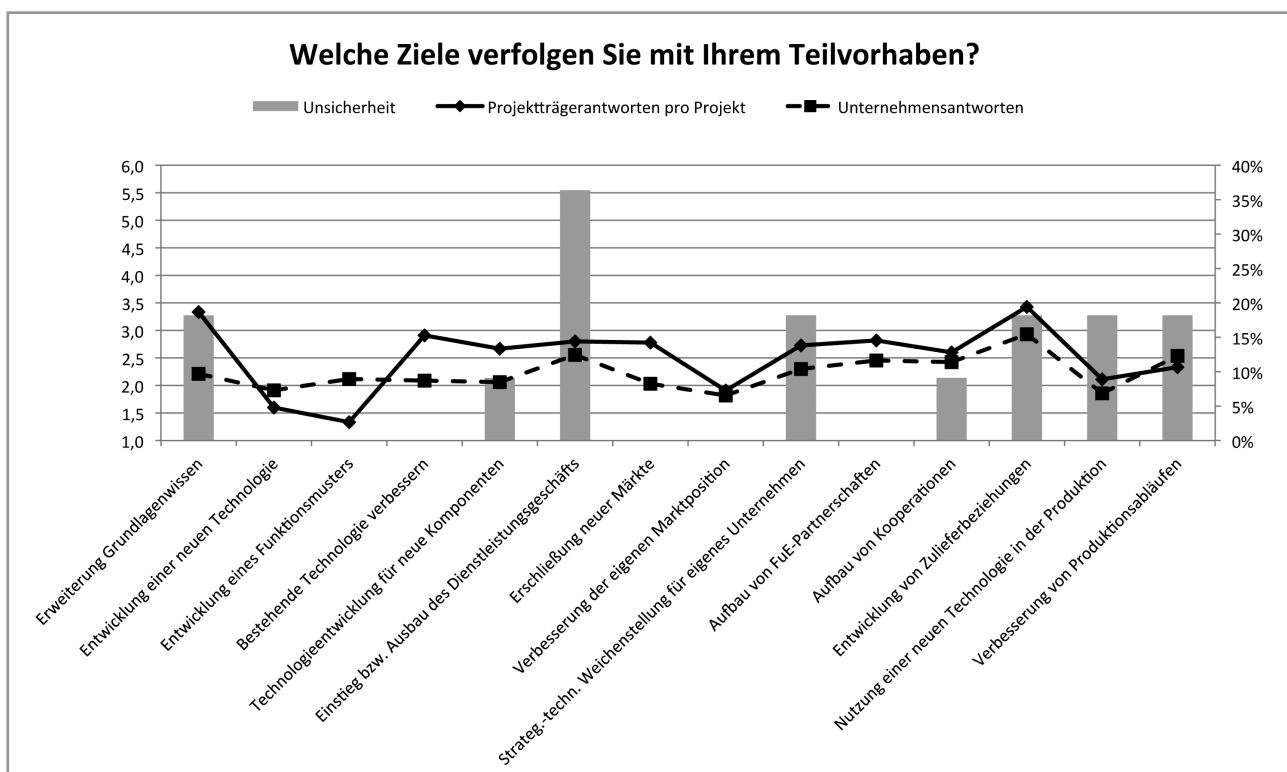


Abbildung 4: Unsicherheit der Projektverantwortlichen bei der Vorhersage von Antworten. Die Prozentzahl gibt an, wie viele Projektverantwortliche sich bei ihrer Einschätzung unsicher fühlten.

UNSICHERHEIT IN BEZUG AUF DIE SIMULATION DES ANTWORTVERHALTENS

Im letzten Abschnitt geht es darum, inwieweit sich die Projektverantwortlichen bei den Vorhersagen des Antwortverhaltens sicher gefühlt haben, also der Auffassung waren, das Antwortverhalten zuverlässig abschätzen zu können und inwieweit das Sicherheitsgefühl mit der realen Abweichung zwischen vorhergesagten und tatsächlichen Antworten zusammenhängt.

Das Ergebnis in Bezug auf die Fragestellung nach den verfolgten Zielen im Teilvorhaben (Abbildung 4) zeigt zunächst anhand der Balken, dass sich die Projektverantwortlichen unterschiedlich sicher in der Einschätzung der verschiedenen Zielstellungen fühlten. Während sie ca. die Hälfte der Aspekte sehr sicher zu beantworten glaubten (kein Balken), besteht bei der anderen Hälfte eine gewisse Unsicherheit bei der Vorhersage der Antworten durch die Unternehmen. Am ehesten unsicher waren die Projektverantwortlichen bei der Einschätzung dazu, wie hoch

die Unternehmen das Ziel „Einstieg bzw. Ausbau des Dienstleistungsgeschäfts“ bewerten würden. Interessanterweise wird hier ersichtlich, dass die gefühlte Unsicherheit bei der Vorhersage nicht im Zusammenhang mit der Treffgenauigkeit steht. Die Übereinstimmung in der Prognose mit den tatsächlichen Antworten war einerseits bei einigen Aspekten sehr hoch, obwohl sich die Projektverantwortlichen eher unsicher fühlten, während andererseits dort die größten Abweichungen auftraten, wo sich die Projektverantwortlichen besonders sicher waren.

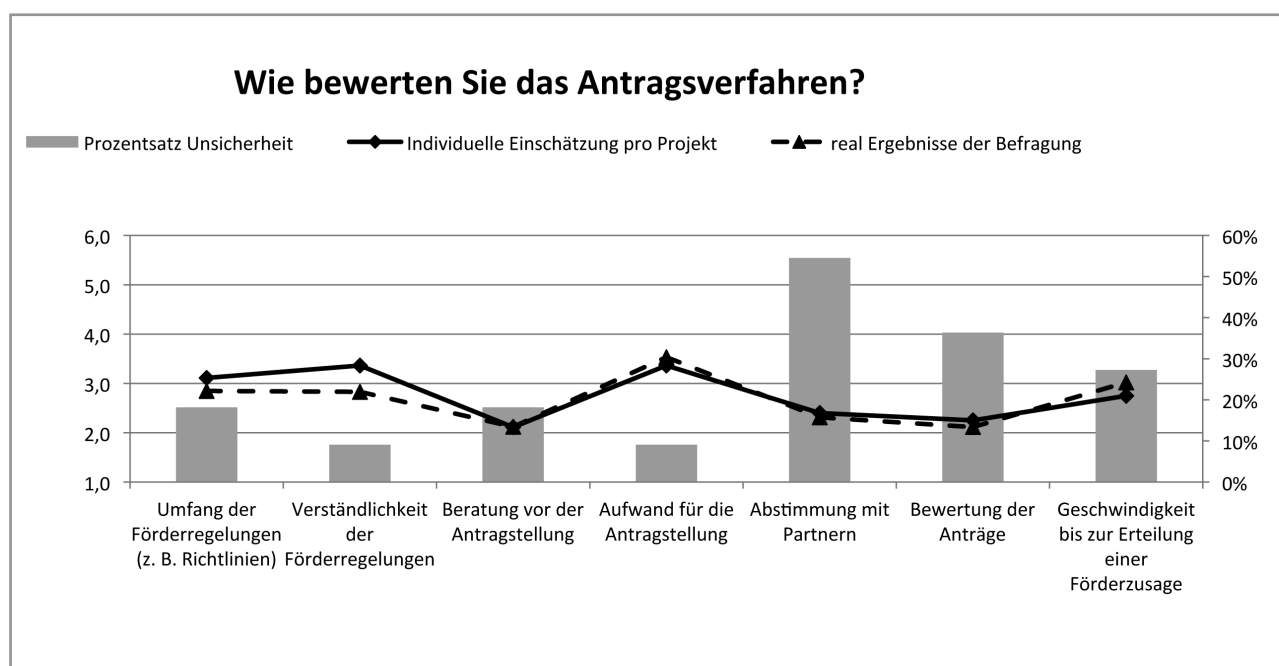


Abbildung 5: Unsicherheit der Projektverantwortlichen bei der Vorhersage von Antworten. Die Prozentzahl gibt an, wie viele Projektverantwortliche sich bei ihrer Einschätzung unsicher fühlten.

Ein ähnliches Bild ergibt sich auch für die gefühlte Unsicherheit bei der Vorhersage der Antworten zum Antragsverhalten (siehe Abbildung 5). Zunächst gab es in Bezug auf die Dimensionen zum Antragsverfahren keine Einzeldimension, bei denen sich alle Projektverantwortlichen sicher fühlten, so wie das bei der Einschätzung zu den Zielen der Fall war. Auch hier ist zu beobachten, dass die Projektverantwortlichen die Bewertung einiger Aspekte des Antragsverfahrens trotz einer gewissen Unsicherheit, was die Unternehmen wohl antworten könnten, sehr genau voraussagen konnten. Im Gegensatz zu einigen Zielen gab es hier jedoch keine Abweichung dahingehend, dass sich die Projektverantwortlichen sehr sicher in Bezug auf das Antwortverhalten waren, es aber zu einer deutlichen Abweichung kam.

Insgesamt konnten die Projektverantwortlichen auch dann, wenn sie sich nicht ganz sicher waren, in vielen Aspekten sehr treffgenau das Antwortverhalten voraussagen. Die gefühlte Unsicherheit bei der Einschätzung des Antwortverhaltens der Projektnehmer ist im vorliegenden Fall also kein guter Gradmesser für die Treffgenauigkeit gewesen.

FAZIT

Als Fazit dieses kleinen Experiments kann festgehalten werden: Nicht jedes Antwortverhalten der Projektnehmer lässt sich durch Projektträger richtig einschätzen. Davon gehen wohl auch die Projektträgermitarbeiter aus, da sie z.T. erhebliche Unsicherheiten hinsichtlich ihrer eigenen Prognosefähigkeit äußerten. Evaluationen erbringen somit durch ihre ergänzenden empirischen Zugänge tatsächlich einen Erkenntnisgewinn, der deutlich über die Einschätzung der Programmverantwortlichen hinausgeht.

Vermutlich ist dieser Erkenntnisgewinn auch in Fragedimensionen, die im hier skizzierten Experiment nicht adressiert wurden, noch deutlich größer. Es ist z.B. zu erwarten, dass Effekte innerhalb der geförderten Institutionen (z.B. Veränderungen in der Arbeitsorganisation oder hinsichtlich der Qualifikationsbasis) für Projektträger nur schwer abzuschätzen sind. Diese Effekte können nur die Projektnehmer selbst bewerten, und dies wiederum wird durch Evaluationen abgefragt.

Wie die unterschiedliche Einschätzung von Projektträger und Fördernehmern zur Frage der Reaktion auf ein Unterbleiben der Förderung zeigt, ist es wünschenswert in Evaluationen nicht nur auf eine Datenquelle für die Beantwortung spezifischer Fragestellungen zurückzugreifen. Die Einschätzung des Projektträgers, dass viele Projekte auch ohne die Förderung durchgeführt worden wären, wirft die Frage auf, warum überhaupt Projekte gefördert werden, bei denen von einer geringen Additionalität ausgegangen wird. Auf der anderen Seite muss gefragt werden, ob die Unternehmen sozial erwünschte Antworten geben. Hier erscheint der Einsatz weiterer Erhebungsmethoden sinnvoll, um eine derartige Diskrepanz aufzudecken. Es wäre interessant zu untersuchen, ob eine systematische Erfassung der Einschätzungen des Projektträgers zur Frage nach der kontrafaktischen Situation Hinweise auf mögliche Mitnahmeeffekte offen legen hilft.

Es gibt zudem Themenkomplexe, bei denen der Versuch einer (ergänzenden) Abschätzung durch Projektträger aus normativen Überlegungen heraus eigentlich nicht unternommen werden sollte. Immer dann, wenn die betroffenen Dimensionen direkt oder indirekt eine Bewertung der Projektträgerarbeit selbst mit umfassen, ist von einem Antwort / Einschätzungsbias auszugehen. Dies gilt für die direkte Bewertung der Projektträgerarbeit, z.B. wie im vorliegenden Beispiel in der Antrags- und

Bewilligungsphase, aber auch für die Beurteilung des Projekterfolgs. Projektträger können zu Recht davon ausgehen, dass die Qualität ihrer eigenen Arbeit auch an der Quote erfolgreicher Projekte gemessen wird. Sie werden diesen Projekterfolg daher möglicherweise systematisch überzeichnen.

Gleichzeitig wird aber auch deutlich, dass Projektträgermitarbeiter „ihre“ Projekte in manchen Dimensionen sehr gut kennen und die zusätzlichen Ergebnisse der Evaluation mit Recht eher als Validierung ihrer eigenen Einschätzungen verstehen. Evaluationsergebnisse bieten in diesen Dimensionen zumindest für den Projektträger keinen echten Erkenntnisgewinn, sondern eher eine offizielle Bestätigung ihrer persönlichen Einschätzungen von unabhängiger Seite.

Offen bleibt aufgrund der begrenzten empirischen Basis des Experiments, in welche Dimensionen spezifische Kenntnisse der Einzelprojekte zu einer „richtigen“ Einschätzung führen, und wo generalisierbare Erfahrungen zum Bewertungsverhalten von Projektnehmern vorliegen, wo also der Projektträger z.B. grundsätzlich und berechtigt davon ausgehen, dass Antragsverfahren für Antragsteller immer als tendenziell zu aufwendig erlebt und bewertet werden. Die Unterschiede zwischen den individuellen Betreuern und dem Gesamtverantwortlichen lassen einerseits darauf schließen, dass spezifische Informationen in einigen Dimensionen für eine hohe Treffsicherheit entscheidend sind. An welchen Stellen eine Einschätzung des Gesamtverantwortlichen zutreffende Ergebnisse liefern kann, da in manchen Dimensionen ggf. längerfristige Erfahrungen zugrunde gelegt werden können, müsste in weiteren Untersuchungen herausgearbeitet werden. Solche Erfahrungswerte sollten zwar prinzipiell auch bei Evaluationseinrichtungen vorliegen, werden aber aufgrund fehlender vergleichender Analysen von Förderprogrammen (siehe Fischl et al. 2013) kaum genutzt.

Auch ist einschränkend anzumerken, dass die im Experiment untersuchten Fragedimensionen ausschließlich Selbsteinschätzungen bzw. Einschätzungen der Projektträgerverantwortlichen zu Selbsteinschätzungen umfassen. Diese müssen im Rahmen einer Evaluation immer um weitere Informationsquellen (z.B. Unternehmensdaten) ergänzt werden, um zu möglichst validen und objektivierbaren Aussagen zu gelangen.

Die Übertragbarkeit der Ergebnisse des Experiments auf weitere Evaluationen bleibt zu prüfen. Möglicherweise gibt es spezifische Maßnahmentypen, in denen die Interaktion zwischen Projektträger und Projektnehmer besonders hoch und damit die Treffgenauigkeit der Einschätzung durch den Projektträger ebenfalls besonders gut wäre. In anderen Maßnahmentypen ist möglicherweise die Akteurslandschaft besonders heterogen und damit eine gute Abschätzung besonders schwierig.

Auch eignet sich die Treffgenauigkeit möglicherweise als Qualitätskriterium für die Arbeit der Projektträger, wenn eine möglichst gute Kenntnis der geförderten Projekte / Projektnehmer als wesentliches Gütemerkmal verstanden wird. In diesem Sinne könnten zukünftig Projektträger selbst externe Evaluationen zum Anlass nehmen, im Sinne eines internen Qualitätsmonitorings begleitende Einschätzungen durch die Projektbetreuer vornehmen zu lassen.

Ein wesentlicher Punkt des Fazits betrifft aber insbesondere die Kooperation zwischen Projektträgern und Evaluatoren. Evaluationen können neue Erkenntnisse für Projektträger generieren, also sollten Evaluationen bei der Interpretation ihrer Daten frühzeitig und eng mit Projektträgern zusammenarbeiten, um ihre Steuerungsinformationen möglichst schnell in den Programmsteuerungsprozess einzubringen. Dabei ist die Unabhängigkeit der Evaluation weiter abzusichern, um normativen Anforderungen ebenfalls gerecht zu werden.

Evaluatoren sollten aber auch die Praxisexpertise der Projektträger nutzen, um empirische Ergebnisse weiter zu validieren und besser zu interpretieren. Mit einer Zunahme an Evaluationen geht einher, dass Projektträger und Evaluatoren häufiger zusammenarbeiten werden. Eine Befragung der Projektträger unter Nutzung des Projektnehmerfragebogens könnte die Mitarbeiterinnen und Mitarbeiter beim Projektträger frühzeitig für spezifische Aspekte sensibilisieren und die Reflexion auch über die eigene Arbeit anregen. Letztlich wird damit auch der Impact der Evaluation selbst erhöht. Die Nutzung der Evaluationsergebnisse erfolgt unserer Meinung nach eher, wenn sich der Projektträger angemessen eingebunden fühlt. Es geht bei Evaluationen direkt oder indirekt letztlich auch um eine Bewertung seiner Leistungen, außerdem sind viele der Empfehlungen ganz konkret vom Projektträger umzusetzen.

In diesem Sinne können die Ergebnisse des Experiments als Ermunterung verstanden werden, eine Evaluationskultur zu fördern, bei der alle verantwortlich Beteiligten auf eine solche Art und Weise eingebunden werden, die dem von ihnen kompetent einzubringenden Wissen entspricht. In einer solchen Evaluationskultur hätte Evaluation vor allem den Charakter eines Reflexions- und Lerninstruments (vgl. Kind / Wessels 2014).

LITERATUR

Astor, Michael/ Fischl, Iris/ Hoffmann, Judith/ Koglin, Gesa/ Kulicke, Marianne/ Sheikh, Sonja/ Wessels, Jan/ Whitelegg, Kathy (2014): Evaluation von Forschungs-, Technologie und Innovationspolitik in Deutschland und Österreich – ein Überblick, in: Böttcher, Wolfgang et. al. (DeGEval-Vorstand) (Hrsg.): Evaluation in Deutschland und Österreich, Stand und Entwicklungsperspektiven in den Arbeitsfeldern der DeGEval – Gesellschaft für Evaluation,

Böttcher, Wolfgang/ Kerlen, Christiane/ Maats, Peter/ Schwab, Oliver/ Sheikh, Sonja (DeGEval-Vorstand) (Hrsg.) (2014): Evaluation in Deutschland und Österreich, Stand und Entwicklungsperspektiven in den Arbeitsfeldern der DeGEval – Gesellschaft für Evaluation. Münster: Waxmann.

Deutsche Bundesregierung: Antwort der Bundesregierung auf die Kleine Anfrage der Abgeordneten Krista Sager, Ekin Deligöz, Katja Dörner, weiterer Abgeordneter und der Fraktion BÜNDNIS 90/DIE GRÜNEN – Drucksache 17/6704 –; Berlin 2011

Fischl, Iris/ Kulicke, Marianne/ Wessels, Jan (2013): Portfolioevaluations: Herausforderungen und aktuelle Erfahrungen. In: Zeitschrift für Evaluation 2/2013, 338-341.

Gigerenzer, Gerd/ Gaissmaier, Wolfgang (2015): Intuition und Führung. Wie gute Entscheidungen entstehen. In: Fröse, Marlies W./Kaudela-Baum, Stephanie/Dievernich, Frank E.P. (Hrsg.), Emotion und Intuition in Führung und Organisation (S. 19-42). Wiesbaden: Springer Gabler.

Kind, Sonja/ Wessels, Jan (2014): Evidenzbasierte Innovationspolitik – Evaluation im Wandel. lit-Perspektive, Working Paper of the Institute for Innovation and Technology - iit | Nr. 18.

AUTORINNEN

JAN WESSELS

Institut für Innovation und Technik

E: wessels@iit-berlin.de

CHRISTIANE KERLEN

Evaluation und Beratung

E: info@kerlen.de

SONJA KIND

Institut für Innovation und Technik

E: kind@iit-berlin.de



openevaluation2016

INTERNATIONAL RTI POLICY EVALUATION CONFERENCE

24 + 25 NOVEMBER 2016

VIENNA | AUSTRIA

www.openevaluation2016.eu

DEAR COLLEAGUE!

Open Evaluation 2016, the largest conference in Europe dedicated to the evaluation of policies in the field of research, technology and innovation policy (RTI) will gather academics, evaluators, research managers, authorities and RTI policy makers to debate challenging developments in RTI policy and their effects on evaluation theory and practice.

The conference takes place in Vienna on 24-25 November 2016. The main conference organisers are the Austrian Platform for Research and Technology Policy Evaluation, the Manchester Institute of Innovation Research and IFRIS - Institut Francilien Recherche Innovation Société, Paris.

The deadline for the call for extended abstracts is 15 April 2016. Please visit our homepage!

The conference addresses new actor settings, approaches and themes in RTI policy evaluation. Thus, the term OPEN EVALUATION signals openness towards new values, new stakeholders and beneficiaries as well as new approaches and themes in RTI policies and RTI evaluations.

THEMES

1 CHANGING THE ROLE OF EVALUATION OF RTI POLICY TO INCORPORATE NEW VALUES, RRI AND NEW FORMS OF SCIENCE

We specifically invite contributions aiming to discuss and assess new values and the changing nature of research, technology development and innovation such as

- Assessments of the effectiveness of new forms of science and innovative approaches (e.g. open digital science)
- Assessment of the relevance and effectiveness of citizen science
- Appraising the inclusion of citizens and civil society organisations in review procedures, in programmatic RTI agenda setting and research processes
- Scientific misconduct within evaluations and identified by evaluations
- Understanding the role of evaluation for responsible research and innovation

2 NEW EVALUATION APPROACHES IN DATA, METHODS, INDICATORS AND THEIR INTERPRETATION

We specifically invite contributions tackling issues such as

- Beyond the ordinary - experiments in RTI evaluations: inspirations, successes and failures
- New developments in evaluation methodologies
- Use of new approaches to data collection and indicator development for evaluation (including altmetrics)
- Societal impacts of agricultural research

3 RTI POLICY EVALUATION IN THE POLICY-MAKING PROCESS

We specifically invite contributions tackling issues such as

- Specific characteristics of RTI as determinants of evaluation approaches
- Organising evaluation for use in RTI policy-making
- Policy learning from evaluations
- Standards and quality assurance mechanisms for safeguarding the societal credibility of RTI policy evaluations
- Enabling conditions for transparent evidence-based RTI policy evaluations
- Connecting research evaluation and public research funding allocations – frameworks, conditions, implications



openevaluation2016

INTERNATIONAL RTI POLICY EVALUATION CONFERENCE

4 SOCIAL INNOVATION, SOCIETAL IMPACT, SOCIETAL CHALLENGES

We specifically invite contributions tackling issues such as

- Evaluation of programmes and instruments supporting social innovations
- Evaluation of societal challenges driven programmes or initiatives
- Measuring the societal impact of RTI policy interventions

5 EVALUATION APPROACHES FOR POLICIES AND PROGRAMMES IN SOCIAL SCIENCES AND HUMANITIES (SSH)

We specifically invite contributions discussing and assessing the use of SSH research inside and outside academia, such as

- Assessment of the impact of SSH research and how it can create benefits in society and economy
- New and alternative metrics for assessing SSH research
- Evaluation of innovative approaches, new pilots and programmes for enhancing and valorising the benefits of SSH in society, economy and policy/government

6 CHALLENGES AND NEW APPROACHES TO EVALUATE EUROPEAN RTI POLICIES AND PROGRAMMES

We specifically invite contributions tackling issues such as

- Evaluation of European (and international) RTI infrastructures
- New approaches in evaluating structural fund investments in RTI
- Evaluation and monitoring of smart specialisation strategies S3
- Evaluation of new approaches and instruments of Horizon 2020 (incl. ERC, FET, JTI, JPI etc.)

Selected papers will be clustered in sessions and panels. Special sessions may complement the programme.

Based on topical fit and quality, selected authors will be invited by the conference organisers to submit their papers to dedicated issues of topical journals.

The extended abstracts of the selected speakers will be published as conference proceedings in the fteval Journal for Research and Technology Policy Evaluation.

AUSTRIAN PLATFORM
for Research and Technology Policy Evaluation



REGISTRATION

www.openevaluation2016.eu

SUBMISSION OF EXTENDED ABSTRACTS

OF 3 TO 5 PAGES INCLUDING REFERENCES UNTIL APRIL 15, 2016

CONFIRMATION ACCEPTANCE: JUNE 30, 2016

CONFERENCE: NOVEMBER 24-25, 2016

MAIN CONTACT: STEFAN PHILIPP // OFFICE@FTEVAL.AT

WE ARE LOOKING FORWARD TO MAKING THIS IMPORTANT EVENT A SUCCESS – TOGETHER WITH YOU.

The Scientific Committee

PREVIOUS ISSUES OF THE FTEVAL JOURNAL FOR RESEARCH AND TECHNOLOGY POLICY EVALUATION:

ISSUE 39

ASSESSING MISSION-ORIENTATED R&D PROGRAMS: COMBINING FORESIGHT AND EVALUATION

Mathias Weber, Wolfgang Polt

ASSESSING THE BROAD SOCIETAL IMPACTS OF RESEARCH: THE CASE OF THE NCCR NORTH-SOUTH PROGRAMME

Claudia Michel, Simon Hearn, Gabriela Wuelser, Thomas Breu

CAN POLICY DESIGN HELP ORGANIZATIONS IMPROVE THEIR NETWORKING CAPABILITIES? AN EMPIRICAL ANALYSIS ON A REGIONAL POLICY

Federica Rossi, Annalisa Caloffi, Margherita Russo

EVALUATION OF R&D INSTITUTIONS IN UKRAINE – THE NEW APPROACH

Vitalii Gryga, Victor Rybachuk, Olha Krasovska

PROMOTING FIRMS' INNOVATIVE BEHAVIOR IN RUSSIA: WHAT WEAKENS THE POWER OF STI POLICY?

Yuri Simachev, Mikhail Kuzyk and Vera Feygina

EVALUATION THROUGH IMPACT: A DIFFERENT VIEWPOINT

Emanuela Reale, Maria Nedeva; Thomas Duncan, Emilia Primeri

MONITORING AND EVALUATION IN JOINT CALLS OF "HORIZONTAL – INCO" ERA-NET AND ERA-NET PLUS ACTIONS

Martin Felix Gajdusek, Nikos Sidiropoulos

NEW MODES OF STAKEHOLDER INVOLVEMENT IN EX-ANTE IMPACT ASSESSMENTS

Susanne Bühner

EVALUATION AS THE CONSTRUCTION OF POLICY NARRATIVES

Erich Prem

WIE WIRKSAM SIND INNOVATIONSFÖRDERMASSNAHMEN IN DER SCHWEIZ?

Franz Barjak

PORTFOLIO EVALUATION: A CASE STUDY TO ILLUSTRATE EVALUATION CHALLENGES

Christiane Kerlen, Jan Wessels, Volker Wiedemer

THE EUROPEAN RESEARCH AREA: A PROCESS-RELATED CHALLENGE FOR INDICATOR AND POLICY DESIGN

Matteo Razzanelli

POSITIONING COMPETENCE CENTRES VIA MONITORING DATA: TOWARDS A SYSTEMATIC APPROACH

Michael Dinges, Jakob Edler, Matthias Weber

MEASURING PRODUCT INNOVATION AND INNOVATIVE CAPACITY: NEW INDICATORS TO EVALUATE RESEARCH PROGRAMMES

Christiane Kerlen, Ernst A. Hartmann

EVALUATION OF CLUSTERS, NETWORKS AND CLUSTER POLICIES – CHALLENGES AND IMPLEMENTATION

Sonja Kind, Gerd Meier zu Köcker

EXPECTATIONS ON THE LONG-TERM IMPACT OF RESEARCH FELLOWSHIPS FROM AN EVALUATION PERSPECTIVE: CHALLENGES AND LIMITS TO MEASURE SIDE-EFFECTS

Christina Schuh

CONFERENCE SESSION SUMMARIES



You can download the fteval Journal for Research and Technology Policy Evaluation from our homepage <http://fteval.at/de/newsletter/archive/>. For a print version please contact office@fteval.at

PREVIOUS ISSUES OF THE FTEVAL JOURNAL FOR RESEARCH AND TECHNOLOGY POLICY EVALUATION:

ISSUE 40

ZWISCHENEVALUIERUNG DES FÖRDERSCHWERPUNKTS
TALENTE

Eva Heckl, Laurenz Wolf

DAS SCHRÖDINGER PROGRAMM DES FWF EVALUIERT.
BEMERKUNGEN VON EVALUATORINNEN UND
AUFTRAGGEBERINNEN

Susanne Bühner, Niclas Meyer, Barbara Zimmermann, Klaus Zinöcker

DER ORGANISIERTE SERENDIPITÄTSEFFEKT ALS
EVALUIERUNGSANSATZ

Fritz Ohler

DIE ROLLE VON BEGLEITENDER EVALUATION UND
BEGLEITFORSCHUNG IN DER FTI-POLITIK. EINE
BESTANDSAUFNAHME.

Peter Kaufmann, Iris Fischl, Sonja Sheikh

BIBLIOMETRIC CASE STUDIES AS COMPLEMENTARY
METHOD FOR S&T PROGRAMME EVALUATIONS

Klaus Schuch, Isabella Wagner, Alexander Degelsegger

WAS KÖNNEN FALLSTUDIEN IM RAHMEN VON
EVALUATIONEN LEISTEN?
ANSPRUCH UND REALITÄT DER ANWENDUNG.

Iris Fischl, Peter Kaufmann

EVALUIERUNG UND VERÄNDERUNG: LERNEN UM ZU
WISSEN, LERNEN UM ZU HANDELN?

Karin Grasenick, Magdalena Pierer, Markus Gruber



EDITORIAL BOARD

Rebecca Allison, **Technopolis UK**; Balázs Borsi, **Eszterházy Károly College**; Elke Dall, **Centre for Social Innovation**; Michael Dinges, **Austrian Institute of Technology**; Leonid Gokhberg, **National Research University Higher School of Economics**; Wolfgang Polt, **Joanneum Research**; Andreas Reinstaller, **WIFO**; Klaus Schuch, **fteval** (chief editor); Michael Stampfer, **WWTF**; Lena Tsipouri, **University of Athens**

AUSTRIAN PLATFORM FOR RESEARCH AND TECHNOLOGY POLICY EVALUATION (fteval)

c/o ZSI – Centre for Social Innovation GmbH

Linke Wienzeile 246, A-1150 Vienna

T +43 1 495 04 42 - 79

F +43 1 495 04 42 - 40

E office@fteval.at

W www.fteval.at

DESIGN

carotte.at

Blindengasse 35/4/97, A-1080 Wien

E krho@subnet.at

PRINT

REMA-Print-Littura Druck- u. Verlags Ges.m.b.H

Neulerchenfelderstraße 35, A-1160 Vienna

AUSTRIAN PLATFORM FOR RESEARCH AND TECHNOLOGY POLICY EVALUATION

Platform fteval members:

Federal Ministry of Science, Research and Economy (bmwfw), Federal Ministry of Transport, Innovation and Technology (bmvit), Austrian Institute of Technology (AIT), AQ Austria, Austria Wirtschaftsservice (AWS), Christian Doppler Research Association (CDG), convelop cooperative knowledge, Austrian Research Promotion Agency (FFG), Austrian Science Fund (FWF), Institute for Advanced Studies (IHS), Austrian Institute for SME Research (KMFA), Joanneum Research, Ludwig Boltzmann Society (LBG), Austrian Institute of Economic Research (WIFO), Austrian Council for Research and Technology Development, Austrian Academy of Science (ÖAW), Technopolis Austria, Vienna Science and Technology Fund (WWTF), Wirtschaftsagentur Wien. Ein Fonds der Stadt Wien, Centre for Social Innovation (ZSI)

AUSTRIAN PLATFORM
for Research and Technology Policy Evaluation



c/o ZSI – Centre for Social Innovation GmbH
Linke Wienzeile 246, A-1150 Vienna

The fteval Journal for Research and Technology Policy Evaluation is a forum to discuss current evaluation practices in the field of research, technology and innovation policy.

T +43 1 495 04 42 - 79
F +43 1 495 04 42 - 40
E office@fteval.at
W www.fteval.at

ZVR-Zahl: 937261837
ISSN-Nr. 1726-6629
©Vienna 2016