

fteval

JOURNAL

for Research and
Technology Policy
Evaluation



ISSUE 40 | DECEMBER 2014

IRIS FISCHL & PETER KAUFMANN

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

PETER KAUFMANN, IRIS FISCHL
& SONJA SHEIKH

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

KARIN GRASENICK, MAGDALENA
PIERER & MARKUS GRUBER

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

FEDERAL MINISTRY FOR SCIENCE,
RESEARCH AND ECONOMY

Minoritenplatz 5, A -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, A-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 Annina Müller-Strassnig

E: annina.mueller@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, A-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, A-1090 Vienna

Dr. Rudolf Novak

E: novak@fwf.ac.at

Dr.ⁱⁿ Dorothea Sturn

E: dorothea.sturn@fwf.ac.at

IHS - INSTITUTE FOR ADVANCED
STUDIES

Stumpergasse 56, 1060 Vienna

Mag. Richard Sellner

E: richard.sellner@ihs.ac.at

JOANNEUM RESEARCH
FORSCHUNGSGESELLSCHAFT MBH

Sensengasse 1, A-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, A-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

TECHNOPOLIS
AUSTRIA

Rudolfsplatz 12/11, A-1010 Vienna

Mag.^a Katharina Warta

E: warta@technopolis-group.com

WIFO - AUSTRIAN INSTITUTE OF
ECONOMIC RESEARCH

Arsenal, Objekt 20, PF 91,
A-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, A-1090 Vienna

Dr. Michael Stampfer

E: michael.stampfer@wwtf.at

Dr. Michael Strassnig

E: michael.strassnig@wwtf.at

ZIT - THE TECHNOLOGY AGENCY OF
THE CITY OF VIENNA

Ebendorferstraße 4/DG,
A-1010 Vienna

Robert Mayer-Unterholzer

E: robert.mayer@zit.co.at

ZSI -
CENTRE FOR SOCIAL INNOVATION

Linke Wienzeile 246, A-1150 Vienna

Dr. Klaus Schuch

E: schuch@zsi.at

Mag.^a Elke Dall

E: dall@zsi.at

INHALT

ISSUE 40 | 2014

- 3 VORWORT
KLAUS SCHUCH
- 5 ZWISCHENEVALUIERUNG DES FÖRDERSCHWERPUNKTS
TALENTE
EVA HECKL, LAURENZ WOLF
- 10 DAS SCHRÖDINGER PROGRAMM DES FWF EVALUIERT.
BEMERKUNGEN VON EVALUATORINNEN UND
AUFTRAGGEBERINNEN
SUSANNE BÜHRER, NICLAS MEYER, BARBARA ZIMMERMANN, KLAUS ZINÖCKER
- 14 DER ORGANISIERTE SERENDIPITÄTSEFFEKT ALS
EVALUIERUNGSANSATZ
FRITZ OHLER
- 20 DIE ROLLE VON BEGLEITENDER EVALUATION UND
BEGLEITFORSCHUNG IN DER FTI-POLITIK. EINE
BESTANDSAUFNAHME.
PETER KAUFMANN, IRIS FISCHL, SONJA SHEIKH
- 27 BIBLIOMETRIC CASE STUDIES AS COMPLEMENTARY
METHOD FOR S&T PROGRAMME EVALUATIONS
KLAUS SCHUCH, ISABELLA WAGNER, ALEXANDER DEGELSEGGER
- 35 WAS KÖNNEN FALLSTUDIEN IM RAHMEN VON
EVALUATIONEN LEISTEN?
ANSPRUCH UND REALITÄT DER ANWENDUNG.
IRIS FISCHL , PETER KAUFMANN
- 43 EVALUIERUNG UND VERÄNDERUNG: LERNEN UM ZU
WISSEN, LERNEN UM ZU HANDELN?
KARIN GRASENICK, MAGDALENA PIERER, MARKUS GRUBER

VORWORT

KLAUS SCHUCH, Geschäftsführer der fteval

LIEBE LESERIN, LIEBER LESER,

Der Fokus der 40. Ausgabe des „fteval- Journal for Research and Technology Policy Evaluation“ liegt diesmal auf methodischen Beiträgen, die durch aktuelle Evaluierungsstudien aus Österreich und Kroatien ergänzt werden.

Im Mittelpunkt der Ausgabe stehen drei längere grundsätzliche Beiträge, welche die fteval auf Wunsch ihrer Mitglieder im Laufe des Jahrs 2014 in Auftrag gegeben hat. Diese beschäftigen sich mit den Anwendungsmöglichkeiten und Begrenzungen von (a) Fallstudien sowie (b) von begleitender Evaluierung für forschungs- und technologiepolitische Evaluierungen. Zusätzlich wird das Thema „Evaluierung und Veränderung“ einer grundsätzlichen Betrachtung unterzogen.

Der Fallstudienbeitrag wurde von Iris Fischl und Peter Kaufmann mit dem Titel „Was können Fallstudien im Rahmen von Evaluationen leisten? Anspruch und Realität der Anwendung“ verfasst. Die beiden Autoren sind gemeinsam mit Sonja Sheikh, alle von der KMU Forschung Austria, auch für den Beitrag über „Die Rolle von begleitender Evaluation/Begleitforschung in der FTI-Politikberatung. Eine Bestandsaufnahme“ verantwortlich, bei dem sie in der Nutzung einen Vergleich zwischen Österreich und Deutschland herstellen. Karin Grasenick, Magdalena Pierrer und Markus Gruber von convelop sind die AutorInnen des Beitrags „Evaluierung und Veränderung: Lernen um zu wissen, lernen um zu handeln?“. Darin positionieren sie Evaluierung im Rahmen des Politikzyklus und diskutieren den Beitrag von Evaluierung zum change management.

Der Beitrag von Isabella Wagner, Alexander Degelsegger und dem Autor dieses Vorworts, alle vom ZSI, über „Bibliometric Case Studies as Complementary Method for S&T Programme Evaluations“ ergänzt den Grundsatzartikel über Fallstudien in der FTI-Evaluierung von Fischl und Kaufmann durch ein praktisches Beispiel, welches im Rahmen der Evaluierung der österreichischen Abkommen zur internationalen wissenschaftlich-technischen Zusammenarbeit angewendet wurde. Er befasst sich mit dem Anwendungsfall bibliometrischer Fallstudien.

Über ein weiteres Beispiel einer kürzlich abgeschlossenen Programmevaluierung in Österreich berichten Susanne Bühner, Niclas Meyer (beide von Fraunhofer ISI), Barbara Zimmermann und Klaus Zinöcker (beide vom FWF). Es handelt sich um die neue Evaluierung des Erwin Schrödinger-Programms des FWF. Darin wird unter anderem ein Abgleich mit den Empfehlungen der vorangegangenen Evaluierung des Programms aus dem Jahr 2006 und insbesondere auch die Erweiterung des Programms durch die Marie Curie Action zur Kofinanzierung nationaler und regionaler Stipendienprogramme thematisiert.

Eva Heckl und Laurenz Wolf, beide von der KMU Forschung Austria, schließlich eröffnen die inhaltlichen Beiträge im vorliegenden fteval-Journal mit einem Ergebnisbericht der Zwischenevaluierung des Förderschwerpunkts Talente der FFG. Dieser umfasst drei Interventionsfelder mit insgesamt acht Programmlinien, für die die AutorInnen differenzierte Schlussfolgerungen entwickeln.

Der Beitrag von Fritz Ohler, Technopolis Group Austria, beruht auf der Evaluation eines komplexen kroatischen FTI-Programms und auf der Evaluation des österreichischen Genforschungsprogramm GEN-AU, wobei sein Schwerpunkt auf dem Phänomen des sogenannten Serendipitätseffekts liegt, also das nicht-intendierte, sondern zufällige Entdecken von Phänomenen "durch einen vorbereiteten Geist". Der Autor thematisiert das Phänomen und seine Behandlung in strikten Evaluierungsdesigns, die auf Basis zumeist rigider ToR entwickelt wurden.

Ich hoffe Sie haben Interesse, Freude und vielleicht sogar Erkenntnisgewinn beim Lesen der genannten Beiträge.

Mein herzlicher Dank geht wie immer an die AutorInnen, an die Reviewer für ihren unermüdlichen Einsatz zur Sicherung der Qualität, sowie an die Graphikerin Caroline Asen und an Daniela Salhofer für ihre umfassende redaktionelle Arbeit.

Ihr

Klaus Schuch

ZWISCHENEVALUIERUNG DES FÖRDERSCHWERPUNKTS TALENTE

EVA HECKL, LAURENZ WOLF

Der vorliegende Artikel beruht auf der Zwischenevaluierung des Förderschwerpunkts Talente für den Zeitraum 2011 bis 2013 und fasst deren Ergebnisse zusammen, wobei sowohl der Förderschwerpunkt als Ganzes als auch die einzelnen Programmlinien untersucht wurden. Der Förderschwerpunkt liegt in der Verantwortung des Bundesministeriums für Verkehr, Innovation und Technologie (BMVIT) und wird von der Österreichischen Forschungsförderungsgesellschaft (FFG) umgesetzt und abgewickelt. Die Zwischenevaluierung wurde von der KMU Forschung Austria im Zeitraum November 2013 bis April 2014 durchgeführt.

ZIELE UND INSTRUMENTE DES FÖRDERSCHWERPUNKTS TALENTE

Der Förderschwerpunkt Talente bündelt alle Aktivitäten zur Förderung der Humanpotenziale im BMVIT (damit wurden Programme wie brainpower austria, FEMtech und generation innovation unter ein Dach zusammengefasst), um die Ausschöpfung des Humanpotenzials im anwendungsorientierten, naturwissenschaftlich-technischen FTI-Bereich zu erhöhen. Der Förderschwerpunkt fokussiert auf den gesamten Karriereverlauf von ForscherInnen im weitesten Sinn, d. h. es sollen (potentielle) ForscherInnen im Kindesalter ebenso gefördert werden wie etablierte ForscherInnen. Der Förderschwerpunkt gliedert sich in drei Interventionsfelder mit insgesamt acht Programmlinien:

- Talente entdecken: Nachwuchs
 - Praktika für Schülerinnen und Schüler
 - Talente regional
- Talente nützen: Chancengleichheit
 - FEMtech Forschungsprojekte
 - FEMtech Karriere
 - FEMtech Praktika für Studentinnen
 - FEMtech Dissertationen
- Talente finden: Forscherinnen und Forscher
 - Karriere-Grants
 - Die österreichische Jobbörse für Forschung, Entwicklung und Innovation

DAS KONZEPT DES FÖRDERSCHWERPUNKTS TALENTE

Bei Betrachtung des Förderschwerpunkts als Ganzes wird ersichtlich, dass dieser alle Karrierestationen (potenzieller) ForscherInnen (Vorschulkinder, SchülerInnen, Studentinnen/DissertantInnen, ForscherInnen) im weitesten Sinn abdeckt, und die verschiedenen Programmlinien ineinandergreifen. Jede „Karrierestation“ (mit Ausnahme der Vorschulkinder) wird von mindestens zwei Programmlinien adressiert. Als Querschnittsmaterie wird „Gender“ zudem in fast allen Programmlinien explizit berücksichtigt (siehe Grafik). Somit entspricht das Konzept von Talente seinem Ziel, Menschen in der angewandten Forschung über den gesamten Karriereverlauf zu unterstützen.

Der Förderschwerpunkt zeichnet sich zudem durch eine hohe Flexibilität und eine große Auswahl an Instrumenten aus, mit denen unterschiedlichste Zielgruppen erreicht werden können und die auch wichtige Schnittstellen und kritische Phasen im Karriereverlauf (Bewusstseinsbildung bei Kleinkindern, Ausbildungsentscheidungen von SchülerInnen, Berufsentscheidungen von Studentinnen, Karriereentscheidungen von ForscherInnen) adressieren. Alle Programmlinien mit Ausnahme von Karriere-Grants setzen nicht nur auf die Förderung des Individuums, sondern auch auf die Verbesserung von Rahmenbedingungen, indem Unternehmen und Forschungseinrichtungen als FördernehmerInnen für die Förderung von Humanpotenzial sensibilisiert werden.

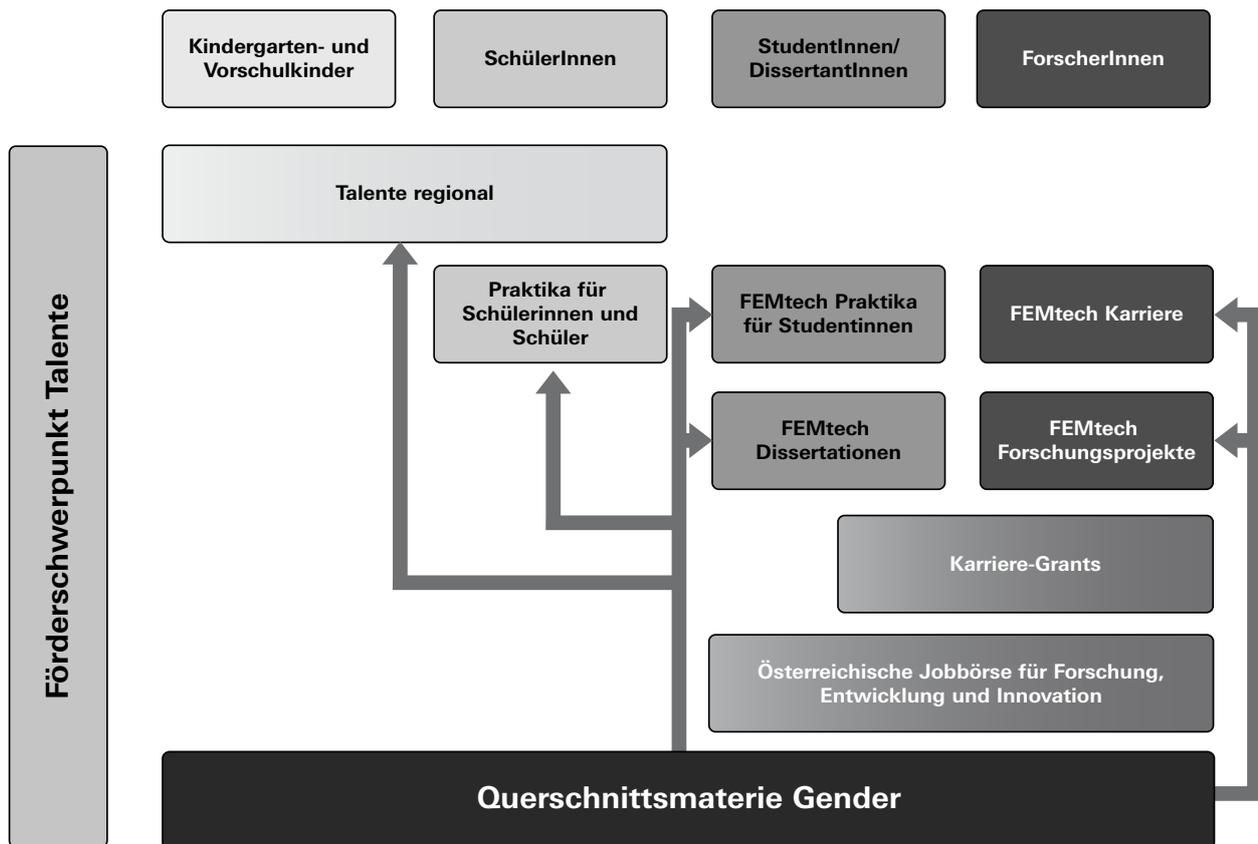
DIE EINZELNEN PROGRAMMLINIEN

Im Folgenden werden die Evaluierungsergebnisse für die einzelnen Programmlinien kurz skizziert. Im Fokus der Analyse standen Fragen, inwieweit die Programmlinien ihre Ziele und ihre Zielgruppen erreichen konnten, und ob sich die Instrumente als adäquat erwiesen haben und von den Zielgruppen angenommen wurden.

Die Programmlinie Praktika für Schülerinnen und Schüler fördert SchülerInnenpraktika in Unternehmen und Forschungseinrichtungen im Zeitraum Juni bis September, um Jugendlichen die Tätigkeiten in der Forschung im naturwissenschaftlich-technischen Bereich näherzubringen. Im Zeitraum 2011 bis 2013 wurden in Summe 4.337 geförderte (Ferial-)Praktikumsplätze für SchülerInnen zur Verfügung gestellt, im Zuge

der letzten Ausschreibung 1.500. Die Praktika wurden mehrheitlich von männlichen Jugendlichen in Anspruch genommen, der Anteil der teilnehmenden Schülerinnen stieg aber kontinuierlich auf zuletzt 38 % im Jahr 2013. Damit konnte die Zielsetzung, ein Drittel der Praktikumsplätze an Mädchen zu vergeben, erreicht und sogar übertroffen werden. Die PraktikantInnen rekrutieren sich zum überwiegenden Teil aus Berufsbildenden Höheren Schulen (BHS) – vorwiegend aus Höheren Technischen Lehranstalten (HTL). Knapp jeder dritte Praktikumsplatz wurde von AHS-SchülerInnen besetzt. Der angestrebte Zielwert von 50 % PraktikantInnen aus Allgemeinbildenden Höheren Schulen (AHS) und sonstigen Schulen

konnte nicht erreicht werden. Dies ist wohl u. a. darauf zurückzuführen, dass in den HTLs verpflichtende Feriapraktika vorgesehen sind, während dies bei AHS-SchülerInnen nicht der Fall ist. Die Verteilung der Praktikumsplätze nach Organisationsart zeigt, dass Unternehmen die deutlich größte Gruppe der PraktikumsgeberInnen bilden, gefolgt von den Universitäten und den außeruniversitären Forschungseinrichtungen.¹



Quelle: KMU Forschung Austria

Die Zufriedenheit der PraktikantInnen liegt auf sehr hohem Niveau und ist im Zeitverlauf nahezu konstant. Fast alle SchülerInnen bewerten das absolvierte Praktikum mit sehr gut oder gut. Der überwiegende Teil der SchülerInnen bewertete die inhaltlichen Anforderungen zudem als (sehr) anspruchsvoll. Mehr als die Hälfte der befragten SchülerInnen gab an, dass ihr Interesse an einem Beruf im technisch-naturwissenschaftlichen Bereich auf Grund des Praktikums gestiegen ist. Auch die PraktikumsanbieterInnen zeigen sich sehr zufrieden. Die Unternehmen und Forschungseinrichtungen ziehen nach eigenen Angaben aus dem Anbieten von Praktika konstant hohen, unmittelbaren bzw. langfristigen Nutzen. Das Praktikumsangebot wird in einem Großteil der Fälle durch die Förderung erst ermöglicht.²

Talente regional fördert kooperative Projekte, die Kindern und Jugendlichen ermöglichen, sich über einen längeren Zeitraum in räumli-

cher Nähe mit FTI-Themen zu beschäftigen. Ein Konsortium muss aus mindestens einem/einer wissenschaftlichen PartnerIn und zwei UnternehmenspartnerInnen mit Innovationsbezug bestehen. Zudem müssen mindestens fünf (vor-)schulische Bildungseinrichtungen aus drei von vier Bildungsstufen beteiligt sein. In den ersten beiden Ausschreibungsperioden wurden 11 bzw. 17 Projekte gefördert. Die ProjektpartnerInnen wollten mit ihren Projekten Kinder und Jugendliche an Naturwissenschaft und Technik heranzuführen, den Kindern und Jugendlichen Berufsmöglichkeiten aufzeigen und ihre (Forschungs-) Tätigkeiten in der jeweiligen Region bekannter machen. Die PädagogInnen sahen v. a. die Möglichkeit, mit wissenschaftlichen Einrichtungen und Unternehmen zu kooperieren. Im Zentrum aller Projekte stand die Begegnung der SchülerInnen und Kleinkinder mit ForscherInnen, meist im Rahmen von Aktionstagen, an denen gemeinsam experimentiert, geforscht, gemessen, analysiert und ausprobiert wurde. Den Abschluss des Projekts bildete in den meisten

1 Quelle: FFG-Monitoring-Daten

2 Quelle: Talente entdecken: Praktika für Schülerinnen und Schüler, Feedback SchülerInnen und Feedback FörderungsnehmerInnen, 2011-2013

Fällen ein Forschungsfest, bei dem die Kinder und Jugendlichen ihre Projekte einem breiteren Publikum präsentieren konnten.

Die interviewten ProjektpartnerInnen sehen die Programmlinie Talente regional nicht nur sehr geeignet, um Kinder und Jugendliche für Naturwissenschaft und Technik zu begeistern, sondern erachten derartige Förderungen auch als absolut notwendig, um das Interesse der Zielgruppe zu steigern. Als Stärke der Programmlinie werden die Offenheit bezüglich der Themen und Herangehensweisen genannt sowie die Langfristigkeit der Projekte. Von großer Relevanz sind auch die hohe Praxisorientierung sowie die zahlreichen Interaktionsmöglichkeiten von ForscherInnen und Kindern/Jugendlichen.³

Im Zentrum der **FEMtech Forschungsprojekte** stehen Forschungsvorhaben aus dem Bereich der angewandten Forschung, die genderrelevante Inhalte abhandeln. In den bisherigen zwei Ausschreibungsrunden wurden jeweils acht Projekte gefördert. Die themenoffene Gestaltung der FEMtech Forschungsprojekte drückt sich in den bisherigen Ausschreibungsrunden durch recht heterogene Anwendungsfelder der geförderten FTI-Projekte aus. Bei den untersuchten FördernehmerInnen handelt es sich überwiegend um Unternehmen und Forschungseinrichtungen, die sich im Zuge ihrer Forschungstätigkeiten bereits mit Genderthemen auseinandergesetzt haben. Für die betrachteten Projekte bzw. die forschenden Organisationen gilt, dass ohne Unterstützung durch die FEMtech Forschungsprojekte die Umsetzung auf Grund unzureichender finanzieller Eigenmittel nicht möglich gewesen wäre. Die spezielle Programmlinie trage dazu bei, die Wahrnehmung unterschiedlicher Lebensrealitäten in oftmals männerdominierte Forschungsfelder einzubeziehen. Die Programmlinie weist eine hohe Zufriedenheit der beteiligten Unternehmen und Forschungseinrichtungen auf und eine neuerliche Teilnahme an FEMtech Forschungsprojekte ist für die Befragten durchaus vorstellbar.⁴

FEMtech Karriere unterstützt strukturelle und nachhaltige Maßnahmen, die zur Chancengleichheit von Frauen und Männern in Unternehmen und außeruniversitären Forschungseinrichtungen im naturwissenschaftlich-technischen Bereich führen. Jedes FEMtech Karriere Projekt besteht aus zwei Pflichtmodulen (Aufbau Genderkompetenz und Projektmanagement) und mehreren Wahlmodulen (aus den Bereichen Öffentlichkeitsarbeit, Personalmanagement, Work-Life-Balance sowie Coaching/Mentoring und Karriereentwicklung). Im Beobachtungszeitraum wurden insgesamt acht FEMtech Karriere Projekte beantragt und gefördert. Damit liegt die Anzahl der geförderten Projekte deutlich unter dem Zielwert von 15 eingereichten bzw. 10 geförderten Projekten.

Bei den FördernehmerInnen handelt es sich Großteils um Organisationen, die sich mit dem Genderthema bereits befasst haben bzw. um Organisationen, deren Personalverantwortliche für das Thema sensibilisiert sind. Eine Mobilisierung von „unbedarften“ Unternehmen und Forschungseinrichtungen scheint nicht gelungen. Der Einsicht über die Notwendigkeit von Personalentwicklungsmaßnahmen im Allgemeinen

und Maßnahmen zur Erhöhung der Chancengleichheit im Besonderen stehen aber die fehlenden Ressourcen gegenüber, daher scheint erst durch die Förderung die Möglichkeit der Durchführung derartiger Maßnahmen gegeben. Auswirkungen der FEMtech Karriere Projekte sind z. B. die Sensibilisierung für die Genderthematik v. a. der Führungskräfte, Änderungen von Recruitingprozessen, eine Erhöhung der Anteile weiblicher TechnikerInnen, die Einführung flexibler Arbeitszeitmodelle, die Beförderung von Frauen in Führungspositionen etc.⁵

Das Förderangebot FEMtech Praktika für Studentinnen zielt auf die Gewinnung von weiblichen Nachwuchskräften im Themenbereich naturwissenschaftlich-technischer Forschungsfelder ab. Es wurden 234 (2012) bzw. 490 (2013) Praktikumsstellen gefördert, in erster Linie in Forschungseinrichtungen (rd. 60 %) und Unternehmen (rd. 30 %).⁶ Die Praktikumsdauer kann je nach Bedarf zwischen einem und sechs Monaten liegen; sechsmonatige Praktika werden am häufigsten vergeben. Im Untersuchungszeitraum hat es zwei Ausschreibungen gegeben, welche trotz zwischenzeitlicher Mittelaufstockung auf Grund von Budgetauschöpfung vorzeitig geschlossen werden mussten.

Für fast jede fünfte Studentin führte laut eigenen Angaben das Praktikum nach Abschluss zu einer weiterführenden Anstellung. Das Feedback der Praktikantinnen ist äußerst gut, sowohl in Hinblick auf die Inhalte als auch die Betreuung während des Praktikums. Durch das Praktikum dürfte auch das Interesse der Studentinnen an naturwissenschaftlich-technischer Forschung gestiegen sein, ebenso wie ihr Fachwissen. Auch die FördernehmerInnen sehen die Praktika sehr positiv. Zwei Drittel der teilnehmenden Unternehmen und Forschungseinrichtungen profitieren dabei nach eigenen Angaben unmittelbar oder langfristig von den Praktikantinnen. Die Additionalitätswirkungen dürften sehr hoch sein, da fast alle FördernehmerInnen angegeben haben, dass sie ohne Förderung keine oder weniger Praktikumsplätze angeboten hätten.⁷

Im Rahmen der Programmlinie **FEMtech Dissertationen bzw. Dissertationen in den Themen des BMVIT** werden Dissertationsprojekte von Personen (erste Ausschreibung nur Frauen – diese wurde näher untersucht) gefördert, die für die Dauer des Dissertationsprojekts in einem Unternehmen bzw. einer außeruniversitären Forschungseinrichtung angestellt werden. Im Beobachtungszeitraum wurden fünf bzw. 13 Projekte gefördert. Die Motivation zur Teilnahme war eher inhaltlicher Natur und weniger der Aspekt der Förderung einer Dissertantin. Durch die Programmlinie gelingt es, Frauen bei der Höherqualifizierung zu unterstützen bzw. diese dazu zu motivieren. Die Dissertantinnen werden von den geförderten Unternehmen und Forschungseinrichtungen sehr umfassend betreut. Insbesondere innerhalb der geförderten Forschungseinrichtungen scheinen die Dissertantinnen dabei auch sehr gut in die jeweiligen Forschungsteams bzw. die Organisation als solches integriert zu sein. Für Unternehmen dürfte dies hingegen mitunter eine große Herausforderung darstellen. Bei drei der ProjektnehmerInnen ist die entsprechende Dissertantin bereits in Vollzeit bzw. zu 5/6 angestellt, bei den anderen beiden ist eine Weiterbeschäftigung der Dissertantin nach

3 Die Analysen stützen sich auf Untersuchung der Endberichte und Fallstudien bereits abgeschlossener Projekte (erste Ausschreibung) und auf eine Online-Befragung von in den Projekten involvierten PädagogInnen (erste und zweite Ausschreibung).

4 Die Analysen basieren auf Fallstudien von Projekten aus der ersten Ausschreibung.

5 Die Analysen erfolgten auf Basis von Fallstudien (fast) abgeschlossener Projekte.

6 Quelle: FFG- Monitoring-Daten

7 Quelle: FEMtech Praktika für Studentinnen: Feedback der Praktikantinnen und FörderungsnehmerInnen, 2011-2013

Projektende geplant bzw. grundsätzlich vorstellbar. Damit dürfte diesen Frauen der Einstieg ins Berufsleben auch durch FEMtech Dissertationen erfolgreich geglückt sein. Von den ProjektnehmerInnen wird die Programmlinie durchwegs sehr positiv beurteilt.⁸

Die Programmlinie **Karriere-Grants** unterstützt ForscherInnen, die eine Beschäftigung am Forschungsstandort Österreich anstreben. Förderbar sind hierbei Reisekosten für Vorstellungsgespräche (Interview Grants), Umzugskosten bei Stellenantritt (Relocation Grants) und die berufliche Integration des/der Partners/Partnerin (Dual Career Grants). Im Beobachtungszeitraum wurden 29 (2011), 130 (2012) bzw. 90 (2013) Karriere-Grants ausbezahlt, vorwiegend Interview und Relocation Grants.⁹

Die **Österreichische Jobbörse für Forschung, Entwicklung und Innovation** ist eine frei zugängliche und kostenlos nutzbare Plattform zur Zusammenführung von JobanbieterInnen und -suchenden im F&E-sowie Innovationsbereich. Ende des Jahres 2013 wurden 1.031 Jobangebote in der Jobbörse angeboten, wobei der Tagesdurchschnitt der angebotenen Stellen bei einem Wert von 333 lag.¹⁰

GESAMTBETRACHTUNG

Wie schon erwähnt, decken die drei Interventionsfelder mit ihren acht Programmlinien alle Karrierestationen im weitesten Sinn ab und die verschiedenen Programmlinien ergänzen einander zu einer „Förderkette“. Von insgesamt knapp 1.200 Organisationen, die Förderungen aus dem Talente-Portfolio in Anspruch genommen haben, sind etwas mehr als die Hälfte Unternehmen, 17 % Universitäten und Fachhochschulen und 8 % außeruniversitäre Forschungseinrichtungen. 12 % der FördernehmerInnen haben zwei oder mehr Programmlinien in Anspruch genommen, und 33 % haben mehrere Förderungen innerhalb einer Programmlinie erhalten. Zudem scheint der Förderschwerpunkt mit seinen teils sehr niederschweligen Angeboten (v. a. Praktika für SchülerInnen) ein „Einstiegsprogramm“ zu sein. 17 % der FördernehmerInnen haben im Rahmen von Talente das erste Mal eine Förderung bei der FFG in Anspruch genommen.¹¹ Die Bedeutung der Genderthematik wird durch die Integration des Programms FEMtech, aber auch durch die Vorgabe, Genderkriterien in anderen Programmlinien zu berücksichtigen, unterstrichen, ebenso wie durch FEMtech Forschungsprojekte, das als einzige Programmlinie nicht die Förderung von Humanressourcen zum Ziel hat, sondern genderrelevante Forschungsprojekte fördert.

SCHLUSSFOLGERUNGEN

Auf Basis der Evaluierungsergebnisse konnten mögliche **Interventionsfelder** in den einzelnen Programmlinien identifiziert werden:

- **Praktika für Schülerinnen und Schüler:** Verstärktes Augenmerk sollte auf SchülerInnen gerichtet werden, deren Weg noch nicht in Richtung eines naturwissenschaftlich-technischen Studiums oder eines Berufs in diesem Bereich vorgezeichnet ist.

Daher sollte die Informationsverbreitung in den AHS verstärkt werden. Die verstärkte Nutzung der Praktikabörse könnte auch zu mehr Ausgewogenheit in Bezug auf den schulischen Hintergrund der PraktikantInnen beitragen.

- **Talente regional:** Aufgrund der besonderen Herausforderungen bei der Koordinierung der vielen ProjektpartnerInnen wird eine Erhöhung des Budgets angeregt bzw. die Möglichkeit dem Projektmanagement mehr Budgetanteil zuordnen zu können. Bei der Einbeziehung der BildungspartnerInnen könnte eine Reduzierung auf zwei Bildungsstufen ein solches Projekt vereinfachen (auch bezüglich der Akquise der BildungspartnerInnen) und zu noch passgenaueren Angeboten führen. Der Zeitpunkt der Ausschreibung bzw. der Förderentscheidung sollte so gewählt werden, dass die Projekte von Beginn an in der Planung des neuen Schuljahres berücksichtigt werden können.
- **FEMtech Forschungsprojekte:** Obwohl es sich hier um klassische Forschungsförderung handelt, hat die Programmlinie insofern ihren Platz im Talente-Portfolio, als von einer ausstrahlenden Wirkung der Inhalte auf das Forschungsfeld und die Organisationen, die dieses Feld bearbeiten, ausgegangen werden kann. Solange genderrelevante Forschungsinhalte nicht in ausreichendem Maße in anderen Programmlinien Berücksichtigung finden, scheint eine derartige Programmlinie im Rahmen von Talente unabdingbar.
- **FEMtech Karriere:** Um einen breiteren Adressatenkreis ansprechen zu können, könnte die Programmlinie adaptiert bzw. gesplittet werden. FEMtech Karriere „light“ könnte erste Maßnahmen zur Implementierung von Chancengleichheit fördern. Hier könnte z. B. nur ein Modul gefördert werden sowie die Förderung (und der Eigenmittelanteil) entsprechend geringer ausfallen. Dies würde die Barrieren für Neueinsteiger bzw. KMU verringern. „Fortgeschrittene“ FördernehmerInnen könnten wie bisher ein volles FEMtech Karriere Projekt beantragen, bei einem Folgeantrag müsste aber gesichert sein, dass neue Elemente mit neuem Fokus ins Projekt integriert werden.
- **FEMtech Praktika für Studentinnen:** Da die Studentinnen durch die einschlägige Studienwahl schon einen Schritt in Richtung Naturwissenschaft/Technik gemacht haben, aber die endgültige Berufswahl noch aussteht, scheinen Faktoren, die die Motivation im Berufsfeld zu bleiben erhöhen, wie dies ein Praktikum darstellt, entscheidend, um dem Phänomen der leaky pipeline (Frauen gehen im Karriereverlauf der Forschung verloren) zu begegnen. Daher sollte eine Aufstockung der Mittel für die FEMtech Praktika für Studentinnen in Erwägung gezogen werden.
- **FEMtech Dissertationen:** Es müsste geprüft werden, in wie weit eine Anhebung der Förderquote für Unternehmen, die erstmals eine Dissertantin / einen Dissertanten aufnehmen bzw.

8 Die Aussagen beziehen sich auf die fünf im Rahmen der ersten Ausschreibung geförderten Projekte und basieren auf Interviews.

9 Quelle: FFG-Monitoring-Daten. Es erfolgte keine nähere Analyse der Programmlinie.

10 Quelle: FFG Quartalsstatistiken Jobbörse. Es erfolgte keine nähere Analyse der Programmlinie.

11 Quelle: FFG-Monitoring-Daten

für KMU sinnvoll bzw. notwendig sein könnte, da diese oftmals über wenig Erfahrung in der Durchführung wissenschaftlicher Projekte verfügen und somit einen deutlich höheren Betreuungs- und Organisationsaufwand haben als Forschungseinrichtungen.

Insgesamt scheint die Bündelung der Aktivitäten des BMVIT im Bereich Humanressourcen gelungen und notwendig, um die Wichtigkeit des Themas für die Forschung an sich zu unterstreichen, denn es gilt, dass es ohne exzellente ForscherInnen auch keine exzellente Forschung geben kann. Zur Sichtbarmachung des Förderschwerpunkts Talente und um eine noch größere Breitenwirksamkeit zu erzeugen, erscheint es wichtig, Talente als Marke stärker zu etablieren und den Förderschwerpunkt in seiner Gesamtheit sowie die einzelnen Programmlinien entsprechend zu promoten.

AUTORINNEN

EVA HECKL

KMU Forschung Austria / Austrian Institute for SME Research

E: e.heckl@kmuforschung.ac.at

LAURENZ WOLF

KMU Forschung Austria / Austrian Institute for SME Research

E: l.wolf@kmuforschung.ac.at

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

SUSANNE BÜHRER, NICLAS MEYER, BARBARA ZIMMERMANN, KLAUS ZINÖCKER

Das Erwin-Schrödinger-Programm wurde 1985 vom österreichischen Wissenschaftsfonds FWF als Auslandsstipendium für junge Postdocs ins Leben gerufen. Im Jahr 2006 wurde das Programm zum ersten Mal mit positivem Befund -, evaluiert: „a globally well performing program“ (Warta 2006, Seite 14); daneben wurde auch Verbesserungspotential identifiziert: Viele Schrödinger-Stipendiatinnen und -Stipendiaten wünschten sich mehr Unterstützung für ihre Rückkehr und – disziplinspezifisch – längere Projektlaufzeiten.

MIT EC-KOFINANZIERUNG PROGRAMM AUSGEBAUT

Mit dem 7. Rahmenprogramm schuf die Europäische Kommission (EC) im ‚People Programme‘ eine Marie Curie Action zur Kofinanzierung

nationaler und regionaler Stipendienprogramme. Die Kriterien für die Mittelvergabe seitens der Europäischen Kommission basierten Großteils auf den Prinzipien der „Charter and Code for Researchers“, die der FWF seit Jahren als Basis seiner Arbeit sieht. Im Wesentlichen sind dies ein transparenter und fairer Auswahlprozess, gute Arbeitsbedingungen für ForscherInnen und ein wesentlicher Beitrag zu deren Karriereentwicklung. Die Kombination der Zielsetzung und der Vergabekriterien zusammen mit dem georteten Bedarf, den Stipendiatinnen und Stipendiaten die Rückkehr zu erleichtern, war ausschlaggebend für den FWF, einen Kofinanzierungsantrag zu stellen. Insgesamt konnte der FWF 13,2 Mio. € für das Schrödinger-Programm einwerben, für dessen Erweiterung die Empfehlungen der Evaluierung (Warta 2006) umgesetzt wurden. Anregungen aus den Gutachten der Europäischen Kommission wurden ebenso in das Design des Programms aufgenommen.

ERWIN-SCHRÖDINGER-AUSLANDSSTIPENDIEN MIT RÜCKKEHRPHASE

ZIELGRUPPE

Hochqualifizierte junge WissenschaftlerInnen aller Fachdisziplinen aus Österreich

ZIELSETZUNG

- Förderung der Mitarbeit an führenden Forschungseinrichtungen im Ausland, Erwerb von Auslandserfahrung in der Postdoc-Phase
- Erleichterung des Zugangs zu neuen Wissenschaftsgebieten, Methoden, Verfahren und Techniken, um - nach der Rückkehr - zur weiteren Entwicklung der Wissenschaften in Österreich beizutragen

ANFORDERUNGEN

- abgeschlossenes Doktorat
- internationale wissenschaftliche Publikationen
- Einladungsschreiben der ausländischen Forschungsstätte
- bei Beantragung einer Rückkehrphase: Bestätigung der inländischen Forschungsstätte

DAUER

10 bis 24 Monate ohne Rückkehrphase bzw. 16 bis 36 Monate mit Rückkehrphase (Rückkehrphase = 6 bis 12 Monate)

HÖHE

Auslandsstipendium: abhängig vom Aufenthaltsort, durchschnittlich 35.000 € pro Jahr steuerfrei;
für die Rückkehrphase: Dienstvertrag mit Senior-Postdoc-Gehalt + 10.000 € p.a.

Die Bewilligungen durch die Europäische Kommission waren für den FWF ein Erfolg in zweifacher Hinsicht: Einerseits hat das Schrödinger-Programm ein „Gütesiegel“ bekommen, in dem Sinne, dass es zur Karriereentwicklung von WissenschaftlerInnen und zur Stärkung des europäischen Forschungsraumes beiträgt, dass das Auswahlverfahren der Stipendiatinnen und Stipendiaten transparent und fair ist und nicht zuletzt, dass das Programm-Management den hohen Ansprüchen der Europäischen Kommission entspricht. Andererseits konnte man dem wissenschaftlichen Nachwuchs ein wesentlich verbessertes Karriereentwicklungsprogramm mit guten Erfolgschancen bieten.

EVALUIERUNGSSTANDARDS UND NEUARTIGER METHODENMIX

Evaluation ist für den FWF ein zentrales Instrument zur Gestaltung seiner Initiativen und Programme. Qualität und Effizienz der Fördermaßnahmen wird von externen, unabhängigen EvaluatorenInnen in regelmäßigen Abständen überprüft. Darüber hinaus bestand die berechtigte Nachfrage seitens der Europäischen Kommission, nach einer fast fünfjährigen Förderungsperiode einen Leistungs- und Wirkungsnachweis einzufordern. Im Rahmen eines Wettbewerbsverfahrens wurde im Sommer 2013 ein Team des Fraunhofer-Instituts für System- und Innovationsforschung ISI unter der Leitung von Susanne Bühner und Niclas Meyer mit einer Evaluierung des Schrödinger-Programmes inklusive seiner Neuerungen beauftragt. Der FWF hatte sich 2012 selbstverpflichtet, externe Programmevaluationen gewissen Standards hinsichtlich Qualität, Transparenz und Controlling zu unterwerfen (vgl. Qualitäts- und Transparenzregeln von FWF-beauftragten Evaluierungen). Ein weiteres wichtiges Element ist die Begleitung dieser Evaluierungen durch internationale ExpertInnen.

Es gibt verschiedene methodische Ansätze, wie Stipendien- und Karriereentwicklungsprogramme auf ihre Wirkung untersucht werden können: eher Interviewbasiert, via Fragebögen, via CV-Tracking, um nur einige zu nennen. Für die Schrödinger-Evaluierung 2013/2014 wählte das Fraunhofer ISI einen innovativen Methodenmix bestehend aus Bibliometrie, Online-Befragung und einem Stakeholder-Workshop. Die bibliometrischen Daten und Methoden wurden dabei nicht bloß zur Erfassung von Publikationszahlen und Zitationsraten eingesetzt. Stattdessen wurden die bibliometrischen Daten hauptsächlich als Adressquelle für die Online-Befragung der StipendiatInnen und die Erstellung einer randomisierten Kontrollgruppe, die ebenfalls online befragt wurde, genutzt. Außerdem wurden diese Daten genutzt um anhand von Kopublikationen die Vernetzung von Wissenschaftlern zu untersuchen.

Als Grundlage für diese Analysen mussten zuerst die SchrödingerstipendiatInnen in der bibliometrischen Datenbank Scopus identifiziert werden. Dies gelang bei 91 % der StipendiatInnen und im Vergleich zur Grundgesamtheit aller StipendiatInnen gab es unter diesen 91 % keinerlei auffälligen Abweichungen im Hinblick auf die generellen soziodemographischen Merkmale der StipendiatInnen.

Ziel des Kontrollgruppenvergleichs, d.h. einer Befragung von zwei Gruppen von WissenschaftlerInnen, die sich lediglich im Hinblick auf das Merkmal „Förderung“ unterscheiden, ansonsten aber vergleichbar

sind (im Hinblick auf die wissenschaftliche Disziplin, das Alter, das Geschlecht), war es, die StipendiatInnen mit Wissenschaftlern mit ähnlichen Startchancen zu vergleichen. Zu diesem Zweck wurden WissenschaftlerInnen aus der bibliometrischen Scopus-Datenbank nach dem Zufallsprinzip ausgewählt, die sowohl das Geschlecht, (Publikations-) Alter, wissenschaftliche Disziplin als auch den österreichischer Hintergrund betreffend, mit den Stipendiatinnen und Stipendiaten übereinstimmten. All diese Merkmale ließen sich aus den bibliometrischen Daten herausfiltern. Weitere Daten wurden über die darauf folgende Online-Befragung ermittelt.

Die Online-Befragung von Geförderten und der Kontrollgruppe verlief planmäßig. Die Rücklaufquoten (40 % bei den Geförderten, 12,5 % bei der Kontrollgruppe) waren, unter Beachtung der Vielzahl von Online-Befragungen zufriedenstellend und erwartbar (niedrigerer Wert bei der Kontrollgruppe). Entscheidend für den problemlosen Ablauf des Kontrollgruppenvergleichs war der Umstand, dass für die Kontrollgruppe dreimal so viele WissenschaftlerInnen zur Teilnahme eingeladen wurden als StipendiatInnen. Im Endeffekt war die Anzahl der Antworten (und die Merkmale der Antwortenden) sehr ähnlich.

Der Survey wurde um eine bibliometrische Analyse zu internationaler Mobilität, Vernetzungseffekten und Publikations-Output ergänzt. Von Bedeutung war darüber hinaus der Workshop mit VertreterInnen aus Politik, Wissenschaft, Administration sowie ausgewählten StipendiatInnen, der es ermöglichte, die Ergebnisse aus Befragung und Bibliometrie zu diskutieren und zu validieren. Die Vorgaben des FWF, einen Methodenmix zur Anwendung zu bringen und eine Triangulation der Ergebnisse zu ermöglichen, wurden somit erfüllt.

ZENTRALE EVALUIERUNGSERGEBNISSE

Die Evaluationsergebnisse zeigen, dass das Schrödinger-Programm in hohem Maße seine Ziele erreicht. Vor allem in Hinblick auf die Karriereentwicklung ist es als ein äußerst erfolgreiches Programm anzusehen. Dass ein Auslandsaufenthalt in der Postdoc-Phase per se karrierefördernd ist, war ein erwartbares Ergebnis, das sich im Ergebnisvergleich mit den WissenschaftlerInnen der Kontrollgruppe widerspiegelt, die keinen Auslandsaufenthalt nachweisen konnten. Der Vergleich der Schrödinger-StipendiatInnen mit der Kontrollgruppe stellt dem Schrödinger-Programm ein herausragendes Zeugnis aus. In diesem Kontext zeigte sich auch, dass Schrödinger-Fellows das Renommée ihrer Gastinstitute im Allgemeinen höher einschätzten, als WissenschaftlerInnen der Kontrollgruppe die auch einen Auslandsaufenthalt realisiert hatten. Die Ziele des Programmes sind höchst kompatibel mit der Motivation der Fellows, ins Ausland zu gehen, die Erwartungen in den Auslandsaufenthalt werden in hohem Maß – in höherem Maße als bei der Kontrollgruppe – erreicht (allen voran: Auslandserfahrung sammeln, neue Methoden und Techniken erlernen, Spezialisierung, Aufbau neuer wissenschaftlicher Kontakte).

Die Wirkung auf die Karriereentwicklung ist beeindruckend: 47 % aller Schrödinger-StipendiatInnen, die vor 2005 ihr Stipendium antraten,

haben heute ordentliche Professuren inne (von allen Fellows mit Projektbeginn vor 1995 sind es sogar 64 %)¹. Damit bestätigt die Evaluation ähnliche Ergebnisse die aus der letzten Evaluation 2006 hervorgingen.

Neben den Wirkungen auf die persönliche Karriereentwicklung untersucht die Evaluierung auch den Einfluss des Programms auf die österreichischen Forschungsstätten bzw. die österreichische Wissenschaft: Die Schrödinger-StipendiatInnen, die in die österreichische Forschungslandschaft zurückkehrten, konnten ihr erworbenes Wissen in hohem Maße nach ihrer Rückkehr umsetzen, indem sie es an der österreichischen Forschungsstätte anwenden und neue Forschungsschwerpunkte mithilfe neuer Methoden und Techniken etablieren konnten. Die bibliometrische Analyse zeigt auch, dass das Schrödinger-Programm die Integration der österreichischen Wissenschaft in die internationale Forschung verbessert. Die Schrödinger-StipendiatInnen publizieren deutlich häufiger zusammen mit internationalen Autoren, als die Wissenschaftler der Kontrollgruppe.

In diesem Zusammenhang ist auch folgendes Untersuchungsergebnis von Bedeutung: Obwohl 33 % der Schrödinger-Fellows nicht innerhalb von zwölf Monaten nach dem Stipendium heimkehrten, fungieren diese Fellows häufig als „Brückenköpfe“, welche die Einbindung in Österreich forschender WissenschaftlerInnen in internationale Netzwerke vorantreiben. Die bibliometrische Analyse von Ko-Publikationen macht deutlich, dass die im Ausland gebliebenen StipendiatInnen weiterhin mit in Österreich ansässigen WissenschaftlerInnen kooperieren und diese in internationale Kopublikationsprojekte einbinden.

Die Fellows zeigen sich in hohem Maße zufrieden mit dem Schrödinger-Programm und allen Programmkomponenten; ebenso findet die Programmabwicklung höchste Zustimmung. Die noch 2006 georteten Schwächen (zu kurze Laufzeit bzw. keine Unterstützung bei der Rückkehr) finden sich nun nicht mehr. Die seit 2009 eingeführte Rückkehrphase scheint bereits zu greifen, auch wenn die Fallzahl der Befragten, die sie bereits in Anspruch nehmen konnten, noch gering ist. Deutlich wurde, dass der FWF mit ihrer Einführung auf einem guten Weg ist: Für 60 % war sie der Hauptgrund, nach Österreich zurückzukehren bzw. für 34 %, überhaupt in der Wissenschaft zu verbleiben.

Mit der Evaluierung 2006, der Umsetzung der darin geäußerten Empfehlungen und der neuerlichen Überprüfung 2013/2014 soll auch eine Einbettung der vom FWF beauftragten externen Evaluierungen in die Maßnahmengestaltung (wenn man so will, in den forschungspolitischen policy cycle) gewährleistet werden, um Verbesserungspotenzial im österreichischen Forschungsumfeld nicht nur zu identifizieren, sondern dieses im Idealfall auch zu heben.

FEHLENDE KARRIEREPERSPEKTIVE IN ÖSTERREICH

So positiv die Bilanz für das Schrödinger-Programm und den FWF ausfällt, so ernüchternd ist teilweise der Blick auf die Motivlage der hoch qualifizierten jungen Forscherinnen und Forscher, sowohl im Ausland zu verbleiben als auch nach Österreich zurückzukehren. Die wenig attraktiven Karriereperspektiven in der Wissenschaft in Österreich wurden sowohl von der Schrödinger-Gruppe als auch den Wissenschaftlern aus der Kontrollgruppe beklagt. In diesem Zusammenhang wurden insbesondere die suboptimalen Bedingungen an den österreichischen Forschungsstätten im Allgemeinen und etwa die Kettenvertragsregelung im Speziellen, so wie die wenig attraktiven Karriereperspektiven in der Wissenschaft in Österreich genannt. Bleiben WissenschaftlerInnen im Ausland, ist hierfür ausschlaggebend, dass sie trotz nachweislich guter Forschungsleistungen keinen Job in Österreich finden können bzw. dass es im Ausland schlichtweg bessere Karriereperspektiven bzw. Forschungsbedingungen gibt. Wie zu erwarten überwiegen bei jenen ForscherInnen, die nach Österreich zurückkehren, private und familiäre Gründe. Die Jobaussichten in Österreich als Motivation für die Rückkehr nach Österreich rangieren unter 10 %.

ZUSAMMENFASSUNG UND FAZIT

Die hohe Zustimmung der Fellows zu allen Programmkomponenten sowie der nachweislich sowohl durch die bibliometrische Analyse als auch die Befragung ersichtliche große Wirkung auf die individuelle Karriereentwicklung und auf die Internationalisierung der österreichischen Wissenschaft zeigt, dass der FWF mit dem Schrödinger-Programm in seiner jetzigen Ausgestaltung jungen WissenschaftlerInnen aller Fachdisziplinen ein zielgruppengerechtes Postdoc-Programm bieten kann, um entscheidende Auslandserfahrung zu sammeln.

Aus der Evaluierung geht klar hervor, dass folgende Elemente wesentlich für eine exzellente wissenschaftliche Karriere sind: Internationale Mobilität; Zeit, sich ausschließlich der Forschung zu widmen; internationale Netzwerke und Kooperationen und das sich daraus ergebende, international sichtbare Publikationsverhalten. All dies bedient das Schrödinger-Programm im höchsten Maße. Das Evaluationsteam schließt daher seinen Bericht mit der dringenden Empfehlung, dass der FWF die Anzahl der geförderten Personen weiter erhöhen soll, da die zunehmende Internationalisierung der Wissenschaft eine immer stärkere internationale Anbindung und Vernetzung der WissenschaftlerInnen verlangt.

¹ Ein direkter Vergleich mit der Kontrollgruppe war in diesem Falle leider nicht möglich, da anders als bei den Geförderten, bei den WissenschaftlerInnen der Kontrollgruppe (insbesondere jenen ohne Auslandsaufenthalt) kein eindeutiger Startpunkt bestimmt werden konnte. Man hätte statt des Zeitraums zwischen Stipendium und Professur, vielleicht das Publikationsalter zum Zeitpunkt der Professur messen und vergleichen können. Aber dann wäre es nicht möglich gewesen an die 2006er Evaluation anzuknüpfen.

Die Evaluation bestätigte den FWF darin, dass das Programm mit den vorgenommenen Adaptierungen auch heute noch – nach fast 30 Jahren – ein wesentlicher Baustein seines Programmportfolios ist und nach wie vor das Instrument für den Start in eine wissenschaftliche Karriere.

LITERATUR

Meyer, N. Bühler, S. (Fraunhofer ISI 2014): Impact Evaluation of the Erwin Schrödinger Fellowships with Return Phase
<http://goo.gl/soA7qC>

Warta, K. (Technopolis 2006): Evaluation of the FWF mobility programs Erwin Schrödinger and Lise Meitner
www.fteval.at/de/evaluation_studies/programmes/2006/441/

FWF: Qualitäts- und Transparenzregeln von FWF-beauftragten Evaluierungen. <http://goo.gl/Cf7HEq>

AUTOR:INNEN

SUSANNE BÜHRER

Fraunhofer Institute for Systems and Innovation Research ISI

E: s.buehrer@isi.fraunhofer.de

NICLAS MEYER

Fraunhofer Institute for Systems and Innovation Research ISI

E: n.meyer@isi.fraunhofer.de

BARBARA ZIMMERMANN

Austrian Science Fund (FWF)

E: barbara.zimmermann@fwf.ac.at

KLAUS ZINÖCKER

Austrian Science Fund (FWF)

E: klaus.zinoecker@fwf.ac.at

DER ORGANISIERTE SERENDIPITÄTSEFFEKT ALS EVALUIERUNGSANSATZ

FRITZ OHLER

KURZFASSUNG

In diesem Beitrag werden zwei abgeschlossene Evaluierungen herangezogen, um einem Phänomen nachzugehen, das in der Wissenschafts- und Innovationsforschung wohl bekannt ist: der so genannte Serendipitätseffekt, also das nicht-intendierte, sondern zufällige Entdecken von Phänomenen „durch einen vorbereiteten Geist“¹. Anhand dieser beiden Evaluierungen wird erstens gezeigt, dass dieser Effekt zu beobachten ist und zweitens, wie er für Evaluierungen verwendet werden kann.

Die eine nimmt Bezug auf die kroatischen, von der Weltbank geförderten Innovationsprogramme. Hier gelang es, jenseits der umfangreichen *Terms of Reference* und der dahinterstehenden Indikatoren den Blick für den Kontext zu schärfen, namentlich für zahlreiche Lernprozesse, sowohl innerhalb der mit der Implementierung betrauten Förderagentur als auch in der nach und nach aufgebauten *Community* von Antragstellern, Fördernehmern, Gutachter/innen, Berater/innen sowie Förderern.

Die andere handelt vom österreichischen Genforschungsprogramm GEN-AU, welches eine Reihe von Einsichten in die Bedeutung von Management und Governance von Universitäten und Forschungseinrichtungen ermöglicht hat. Hier zeigt sich eine starke Präsenz des Matthäus-Prinzips² sowie eine damit verbundenen Tendenz zu einer *institutional divide* innerhalb der österreichischen Forschungslandschaft.

Gemeinsam ist beiden Evaluierungen – im Nachhinein wenig überraschend – die Rolle des Kontexts und im Speziellen die Schlüsselrolle von Management und Governance, bisweilen auch von handelnden Personen und Persönlichkeiten als spezifischer Kontext. Damit wird die übliche Wirkungskette Input → Output → Outcome → Impact um eine wichtige Variable reicher, gleichzeitig werden das Konzept der Wirkungskette und damit auch manche Interventionsmodelle fragwürdig, wohingegen Kontexte immer eine maßgebliche Rolle spielen.

DAS WELTBANK-PROGRAMM IN KROATIEN: DIE AUSBLENDUNG VON KONTEXTEN DURCH BEST PRACTICES

DAS JUNGE KROATIEN: MODERNISIERUNG, AUFBAU VON INSTITUTIONEN, NUTZUNG DER CHANCEN DURCH NEUBEGINN, UNTERSTÜTZUNG DURCH DIE WELTBANK

2002: Die Republik Kroatien ist gerade zehn Jahre alt, der Kroatienkrieg (1991-1995) erst ein paar Jahre vorbei, Wirtschaft und Gesellschaft müssen erneuert, Institutionen auf- und umgebaut werden. Forschungs- und Innovationspolitik gehören zu diesen bevorzugten Erneuerungsaufgaben. Die Weltbank ist zur Stelle und tut das, was in solchen Situationen pragmatischer Weise getan werden kann: Sie legt Programme auf, also sachlich und zeitlich begrenzte, fokussierte Aktionen, die an wohlausgesuchten Stellen des Systems ansetzen.

Sie findet vor allem eine umfangreiche öffentliche Forschungsstruktur vor, die Universitäten sowie eine – für die ehemals sozialistischen Länder recht typisch – üppige Ausstattung mit öffentlichen Forschungseinrichtungen. Nichts liegt also näher, als diese durch geeignete Maßnahmen dazu zu bringen, ihren Beitrag zur Modernisierung von Wirtschaft und Gesellschaft zu leisten. Die maßgeblichen Strategien und zugehörigen Programme sind³: (i) Gründung von wissensbasierten Unternehmen (RAZUM), (ii) Kooperation von Universitäten und Forschungsorganisationen mit KMU (IRCRO), (iii) Aufbau von Technologie(transfere)zentren und technologischen Infrastrukturen an Universitäten und öffentlichen Forschungseinrichtungen (TEHCRO), (iv) Transfer und Kommerzialisierung von Forschungsergebnissen aus Universitäten und öffentlichen Forschungseinrichtungen (TEST).

Die Programme wurden in 2002 eingeführt und nach fünf Jahren um weitere fünf Jahre verlängert. Vorbilder für diese Art von Programmen gab es zuhause – so genannte *best practices* – und man erkennt in ihnen den Geist der 1990er Jahre. Im Einzelnen wurden solche *best practices* aus dem benachbarten Triest, aus Israel und aus Deutschland importiert.

1 Robert K. Merton: *Social Theory and Social Structure*. The Free Press, Glencoe, Ill. 1957. Der "vorbereitete Geist" findet hier auf Seite 12 folgende Definition: „the discovery through chance by a theoretically prepared mind of valid findings which were not sought for“. Der Serendipitätseffekt hat neuerdings in Form des Surfens im Internet ungeahnte Bedeutung erhalten. Nicht ganz zufällig werden die entsprechenden Suchmaschinen auch Browser genannt (engl. browse around: sich umsehen, stöbern).

2 Robert K. Merton (1985), Robert K. Merton: Der Matthäus-Effekt in der Wissenschaft. In: R. K. Merton (Hrsg.): *Entwicklung und Wandel von Forschungsinteressen*. Suhrkamp, Frankfurt 1985. Sein Name leitet sich aus dem Matthäusevangelium ab: „Denn wer da hat, dem wird gegeben, dass er die Fülle habe; wer aber nicht hat, dem wird auch das genommen, was er hat.“ oder auch, etwas profaner: "success breeds success".

3 <http://www.bicr.hr/en/programs>

Wie sich rückblickend herausstellt, waren die von der Weltbank finanzierten Programme mit insgesamt ca. 35 Mio. Euro für eine Laufzeit von fünf Jahren die größten unter den in Kroatien durchgeführten FTI-Programmen und sie sollten, im Falle positiver Ergebnisse, in eine dritte Runde gehen.

Sie standen wegen des *best practice* Status, wegen des Volumens und wegen der besonderen Präsenz der Weltbank unter besonderer Beobachtung – durch das zuständige Ministerium, durch die Förderagentur BICRO (*Business Innovation Center of Croatia*), nicht zuletzt durch die Weltbank selbst.

DIE TERMS OF REFERENCE FÜR DIE EVALUIERUNG DER WELTBANK-PROGRAMME LAUFEN GEFAHR, DIE FALSCHEN FRAGEN ZU STELLEN

Als die Weltbank am Ende der zweiten Förderperiode 2013 die Evaluierung dieser Programme in Auftrag gibt, werden große Batterien aufgefahren: 27 allgemeine Fragen für jedes der vier Programme sowie 5-7 programmspezifische Fragen, also über 30 Fragen pro Programm. Auch die Liste der Output-, Outcome- und Impactindikatoren ist beeindruckend, ihre Zahl rangiert zwischen 10 und 33 für jedes Programm. Und das zugrunde liegende Interventionsmodell ist ebenso einfach wie direkt: Man schiebt Geld für gute Projekte hinein und man erwartet sich Outputs (Produkte und Dienstleistungen), die dann zu Impacts führen, also Firmengründungen, Umsätze, Arbeitsplätze, Steuern.

Das für das gesamte Kroatien kurz zuvor durchgeführte OECD Innovation Policy Review⁴ zeigt indes ein deutlich anderes und im Einzelnen viel weniger optimistisches Bild als jenes, das in den Weltbankprogrammen vermittelt wird: Die öffentlichen, mit Forschung befassten Institutionen (Universitäten, Forschungseinrichtungen) sind nicht auf Technologietransfer, interdisziplinäre Forschung, Patentverwertung, Ausgründung etc. eingestellt, sondern warten auf die überfällige Reform des Hochschulsektors – die sie übrigens, als es soweit ist, selbst verhindern. Zugleich gibt es zu viele zu kleine Förderprogramme aus nationalen Quellen mit durchwegs zu kurzer Laufzeit, also insgesamt Programme mit geringer Bodenhaftung und Nachhaltigkeit; die meisten sind angebotsseitig auf Forschung und deren Verwertung ausgerichtet. Die umfangreichen, mit der Europäischen Kommission im Zuge der Vorbereitungen des EU-Beitritts Kroatiens verhandelten Strukturreformen und Regulierungen werden von den mit Forschungs- und Innovationspolitik befassten Akteuren nicht als Gelegenheit für Modernisierung aufgegriffen, sondern als Belastung empfunden. Die überwiegend angebotsseitige Innovationsförderung hat zur Folge, dass die Innovatoren häufig auf ihren Projekten sitzenbleiben, weil andere, bisweilen wichtigere Unterstützungsmaßnahmen fehlen – allen voran die Unterstützung beim Betreten internationaler Märkte und damit die Klärung der Frage, ob es etwa besser wäre, einen internationalen Vertrieb aufzubauen, eine Partnerschaft einzugehen oder die Firma an ein international tätiges Unternehmen zu verkaufen. Hier fehlen sowohl das Bewusstsein als auch die Expertise, um solche Prozesse zu unterstützen.⁵

Mit den von der Weltbank finanzierten Programmen wurde also Unterstützung für etwas geboten, wo es – verglichen mit anderen Schwierigkeiten – gar keinen großen Engpass gab, nämlich findige und umtriebige Leute in und im Umfeld von Universitäten und öffentlichen Forschungseinrichtungen anzusprechen und in ihren Innovationsbemühungen zu unterstützen. Umgekehrt wurde von den Universitäten und Forschungseinrichtungen in einer Weise Gebrauch gemacht, auf die diese – und nicht nur die kroatischen – strukturell und aufgrund ihrer Mission nicht vorbereitet waren und sind, nämlich Forschungsergebnisse in Produkte und Dienstleistungen zu verwandeln und diese auf den Markt zu tragen. Forschung in diesen Institutionen wird bekanntermaßen von anderen Motiven und Kalkülen getrieben: Ausschachten einer Theorie, einer Methode, einer Laborkapazität, guter Zugang zum Forschungsobjekt. Die bevorzugten Ergebnisse sind Publikationen, Mitgliedschaft in und Anerkennung durch einschlägige *Communities*, persönliches Fortkommen, Reputation, ein kleines Zubrot hier und dort.

Zusammenfassend: Die von der Weltbank geförderten Programme haben, obwohl auf die Wirtschaft ausgerichtet, einen starken Angebotsbias, die Universitäten und öffentlichen Forschungseinrichtungen werden in ihrer Rolle verkannt, und im Besonderen überschätzt, Unzulänglichkeiten im Institutionensystem werden nicht angesprochen, dem institutionellen Lernen auf Agentur- und Politikebene wird keine Aufmerksamkeit geschenkt und schließlich ist Forschung nicht der einzige, und vielfach nicht einmal der wichtigste Engpass bzw. Risikofaktor.

DIE HAUPTZIELE DER FÖRDERPROGRAMME WERDEN ZWAR NUR BEDINGT ERREICHT, ABER ES GIBT GROSSE ERTRÄGE BEI EINIGEN NICHT-INTENDIERTEN ZIELEN

Funktioniert haben die Förderprogramme im Wesentlichen dort, wo nach und nach und vor allem von der mit der Umsetzung betrauten Agentur maßgebliche Zusatzaktivitäten geleistet wurden: Erstellung von Businessplänen als größerer konzeptueller Rahmen für Geschäftstätigkeiten und in der Folge für Förderansuchen und Förderentscheidungen, Management- und technischer Support im Zuge der Umsetzung der Projekte durch Beiziehen von externen Berater/innen, regelmäßige Reviews durch Mitarbeiter/innen der Förderagentur bzw. durch Einbeziehen Dritter. Diese Zusatzaktivitäten und das aktive Eingreifen in die Durchführung der Projekte durch die Förderagentur BICRO waren in den Programmen nicht geplant, sondern haben sich erst im Zuge der Implementierung konkretisiert und etabliert. Tatsächlich hat BICRO im Lauf der Zeit ein Netzwerk von "helfenden Händen" in Form von externen Gutachter/inne/n und Berater/inne/n aufgebaut und im eigenen Haus die Fähigkeit entwickelt, dieses Netzwerk zu managen, und, nicht zuletzt, zu entscheiden, welche Fragen eher in-house und welche besser mit externen Kräften erledigt werden können und sollen.

Stellt man also die Frage nach den wichtigsten und nachhaltigsten Ergebnissen und Wirkungen der Weltbank-Programme, so zeigen sich recht bescheidene Ergebnisse, was die üblichen Wirkungsindikatoren betrifft: Patente, Firmengründungen, neue Produkte und Dienstleistungen.

4 OECD (2013), *Review of Innovation Policy: Croatia*

5 Fritz Ohler (2014), *Ex Post Evaluation of BICRO's Technology Programmes (PPA-CS-XX)*. Final Report to the Ministry of Science, Education and Sports of the Republic of Croatia, February 2014

gen, Beschäftigte, Erhöhung der Produktivität und der Wettbewerbsfähigkeit der kroatischen Wirtschaft – auch das stand auf der Liste der Terms of Reference.⁶

Verschiebt man jedoch die Perspektive und blickt nicht nur auf die Wirkungen in der kroatischen Wirtschaft, sondern auch auf das Politik- und Institutionensystem, so zeigen sich durchaus substantielle Wirkungen der Förderprogramme der Weltbank. Die eine Gruppe von Wirkungen kreist um die mit der Abwicklung befasste Förderagentur BICRO. Die bloße Existenz der Programme über einen längeren Zeitraum hatte zur Folge, dass die Agentur eine Aufgabe und Herausforderung erhalten hat, an der sie sich abarbeiten und ihr Handwerk lernen konnte. Angesichts des für Kroatien recht typischen Eröffnens und Schließens von Agenturen und angesichts des häufigen Wechsels der Schlüsselakteure in den Ministerien (Minister/innen bzw. Vizeminister/innen, deren Berater/innen, leitende Mitarbeiter/innen) und der internen Strukturen haben die Weltbank-Programme wesentlich zur Stabilisierung von BICRO als Agentur beigetragen, die nunmehr ihr Handwerk versteht und in der Folge einen stabilisierenden Faktor in der ansonsten fragilen Institutionenlandschaft darstellt.

Die zweite Gruppe von Wirkungen betrifft abermals das Institutionensystem und dessen Lernen. Die Existenz der Programme über rund ein Jahrzehnt und die damit verbundene kontinuierliche Befassung hat zum Aufbau und zur Entwicklung einer Community von Leuten aus Universitäten, Forschungseinrichtungen und Firmen, Gutachter/inne/n und Berater/inne/n aus dem In- und Ausland sowie der Förderagentur selbst beigetragen. Diese Community hat ihrerseits – in unterschiedlichen Rollen – das Geschäft der Beantragung, Begutachtung, Abwicklung von Projekten und in weiterer Folge der Unterstützung durch die Förderagentur oder durch Dritte gelernt. Die Förderagentur BICRO weiß inzwischen, wo und wann sie die Rolle des Polizisten, des Arztes oder des Coaches annehmen muss und ob sie dies bevorzugt mit eigenen Kapazitäten tut oder eher Dritte einbezieht.

NICHT-INTENDIERTE NEBENEFFEKTE: DER SERENDIPITÄTSEFFEKT IN DER FTI-FÖRDERUNG UND IHRE ENTDECKUNG IM ZUGE DER EVALUIERUNG

Daraus lässt sich einiges lernen: Die Orientierung an *best practices* blendet meist den Kontext aus, aus dem heraus das Attribut best begründet ist. Obwohl konzeptuell längst überholt, dominiert das lineare Innovationsmodell immer noch die Politikgestaltung: Geld bzw. Kostenreduktion für ein spezifisches F&E-Projekt hinein, Outputs und Impacts aus dem geförderten Projekt heraus. Schließlich verstellt die Konzentration auf Projekte den Blick auf die Förderung der größeren organisatorischen Einheiten mit **Hilfe von Projekten**, also des Unternehmens, des Instituts oder des Teams. Die Folge ist, dass Projekte mehr nach projektspezifischen Kriterien eingeschätzt werden als durch die Frage nach attraktiven und glaubwürdigen Entwicklungslinien der jeweiligen Organisation und die Frage, ob die Projekte *fit-for-purpose* sind.

Die Wahl der Evaluierungsmethoden ist ausgesprochen sensibel für das Evaluierungsergebnis: Die weit über 100 Fragen und die einschlägi-

gen quantitativen Indikatoren verführen zur Verwendung von Fragebögen. Diese sind nicht immer verlässlich, vor allem gibt es einen weithin unkontrollierbaren kulturellen Bias, vor allem, wenn es um Einschätzungen geht. Ferner wird auf diese Weise in der Regel auch in der Evaluierung der Kontext ausgeblendet, namentlich der Kontext der Projekte, der Firmen bzw. Institute, nicht zuletzt der politischen Akteure, vor allem der Agentur selbst. Schließlich trübt der Zwang, die einzelnen und meist umfangreichen *Terms of Reference* auf den Buchstaben genau zu erfüllen, den Blick der Evaluator/inn/en in jene Bereiche, die nicht in den *Terms of Reference* angeführt sind. Angesichts des üblichen Projektdrucks will man sich nicht auch weitere zu den ohnedies schon zu vielen Fragen aufladen.

Die Methoden, die für die Evaluierung der Weltbank-Programme verwendet wurden, waren jede für sich unspektakulär: Auswertung von Programmdokumenten und Monitoring-Daten, Interviews mit Schlüsselpersonen, ein kleiner Fragebogen, in denen Fakten, aber keine Einschätzungen abgefragt wurden, hauptsächlich zur Überprüfung der Konsistenz und Vervollständigung von Indikatoren, die von der Agentur im Zuge ihres Monitorings gesammelt wurden. Die entscheidende Methode waren aber Workshops, in denen sowohl Fördernehmer als auch Fördergeber teilnahmen und in denen die jeweils geförderten Projekte präsentiert und diskutiert wurden. Ziel dieser Workshops war es, gemeinsam zu lernen, vor allem durch Kontextualisieren der Projekte im geförderten Unternehmen / Institut, der Förderfälle im Programm und beide im weiteren Kontext. Dieser umfasst das Wettbewerbsumfeld, Kooperationen, verfügbare und knappe Umfeldfaktoren, Schwierigkeiten, Fachkräfte zu finden, Finanzierung der nächsten Schritte, Schwierigkeiten, den Exportmarkt zu betreten, oder schlicht und einfach, den oder die nächsten Schritte zu absolvieren. Weil Fördernehmer, Fördergeber und Evaluator/inn/en am selben Tisch saßen und gemeinsam diskutierten, gab es eine Fülle von Beobachtungen und Lernmöglichkeiten, vor allem durch unmittelbares Feedback nach jeder Diskussionsrunde bzw. nach jedem Tag.

DER KONTEXT ALS BEVORZUGTE QUELLE VON SERENDIPITÄTSENTDECKUNGEN

Die Serendipitätsentdeckung, die in diesem Evaluierungsprojekt gemacht werden konnte, ist die zentrale Rolle von Kontexten: (i) des Projekts, weil dieses eingebettet ist in andere (Förder-)Projekte, (ii) des Förderempfängers, weil auch noch andere Tätigkeiten und Fähigkeiten erforderlich sind, damit Innovationen gelingen, (iii) des Programmes selbst, insofern gewisse Aktivitäten gefördert werden, andere aber nicht, und dadurch zwangsläufig selektive Aufmerksamkeit entsteht, die von der Agentur verstärkt, aber auch kompensiert werden kann, (iv) schließlich der Förderagentur, ihres Portfolios und ihres institutionellen Umfeldes.

Da "der Zufall nur einen vorbereiteten Geist begünstigt" sind die Evaluatoren und ihre Auftraggeber gut beraten, dem Kontext eine weit höhere Aufmerksamkeit zu schenken als üblich, und diesen Fragen aktiv und kreativ nachzugehen.

6

Eine aus den über hundert Evaluierungsfragen etwa lautete: "Could quantitative results be obtained in terms of: number of start-ups, number of licenses, number of patents, number of jobs created within the Universities and SMEs (during programme implementation), number of jobs created by companies (during and after programme implementation), number of jobs created in companies (after programme implementation), money provided by the programme, capital attracted by companies / spin-outs ...?"

DAS ÖSTERREICHISCHE GENFORSCHUNGSPROGRAMM GEN-AU

GEN-AU: EIN PROGRAMM MIT HOCHFLIEGENDEN AMBITIONEN

Das GEN-AU-Programm⁷ des jetzigen Bundesministeriums für Wissenschaft, Forschung und Wirtschaft ist eines der größten seiner Art: 85 Mio. Euro für eine thematische Forschungsförderung über einen Zeitraum von etwa zehn Jahren, gestartet im Jahr 2001. Die Ambitionen waren groß und umfassend, und spiegeln die damalige Aufbruchsstimmung in der Forschungspolitik wider: die Forschung voranbringen, wirtschaftliche Erträge erzielen, Fachleute ausbilden, gesundheitspolitische Akzente setzen, und, nicht zuletzt, durch sozialwissenschaftliche Analysen zur Debatte über politisch-ethische Fragen der Genforschung beitragen. Auch die Vielfalt der Projektformate war ohne Beispiel: Verbundprojekte, Netzwerkprojekte, Pilotprojekte, transnationale Projekte, ELSA-Projekte⁸ assoziierte Projekte.

Auf Seiten der Förderempfänger gab es für eine Handvoll von Forschungseinrichtungen eine üppige Ernte. Sie warben häufig mehrere, aufeinander abgestimmte größere Projekte ein und waren, auch begünstigt durch einschlägige Förderbedingungen, dadurch in der Lage, neue Forschungslinien zu eröffnen und diese zu entfalten. Dabei zeigen sich einige bemerkenswerte Muster: Erfolg in der Förderung und Erfolg in der Forschung korrelieren mit effizientem Management. Dieses wiederum zeichnet sich durch einige wenige Faktoren aus: interne Qualitätssicherung, etwa dadurch, dass Förderanträge bereits Wochen vor Einreichschluss fertig sind und einer Vorbegutachtung unterzogen werden; enge Abstimmung der Forschung mit Investitionen in Forschungsinfrastruktur und Zugang zu dieser; hohe Aufmerksamkeit für Personalfragen (Rekrutierung, Karriereentwicklung). Nicht, dass GEN-AU diese hohe Kultur des Forschungsmanagements hervorgebracht hat. Vielmehr hat GEN-AU zur Entwicklung und zum Ausbau dieser Kultur beigetragen, wie dies in ähnlicher Form nunmehr durch ERC-Grants geschieht, oder seit längerem durch Spezialforschungsbereiche oder Doktoratskollegs – sie steigern die schon vorgezeigte Leistungsfähigkeit der Institute bzw. der Schlüsselpersonen und ihren Teams.

DIE GEN-AU-EVALUIERUNG ALS REICHE QUELLE VON EINSICHTEN IN MANAGEMENT UND GOVERNANCE VON FORSCHUNG: DER MATTHÄUS-EFFEKT UND DIE BEDEUTUNG DES STANDORTS, VON CORE FACILITIES UND DES FORSCHUNGSMANAGEMENTS

Tatsächlich hat der systematische Blick auf zehn Jahre GEN-AU⁹ und die jeweils großen Förderempfänger – und damit auch einen Großteil der eingesetzten Fördermittel – Licht in einige Grauzonen gebracht, und zwar nicht nur, ja nicht einmal primär von GEN-AU, sondern der öffentlichen Forschung generell. Und dies sind die wichtigsten Beobachtungen.

Das Vorliegen des Matthäus-Effekts. Eines der wichtigsten beobachteten Phänomene ist das Vorhandensein des so genannten Matthäus-Effekts. Er zeigt sich vor allem darin, dass dort, wo größere GEN-AU-Projekte eingeworben wurden, sich häufig auch andere "große Brocken" im Portfolio finden (Spezialforschungsbereiche, Doktoratskollegs, in jüngerer Zeit ERC Grants, bisweilen Wittgenstein-Preise). Diese Agglomerationen sind besonders auffällig bei den eher jüngeren und privatwirtschaftlich organisierten Instituten der ÖAW (*Center of Molecular Medicine (CeMM)*, *Gregor Mendel Institute (GMI)*, *Institute of Molecular Biotechnology (IMBA)*), den *Max F Perutz Laboratories (MFPL)* sowie dem *Institute of Molecular Pathology (IMP)*, das als privates Unternehmen im engeren Sinn offenbar in der Lage ist, in den vorderen Reihen der akademischen Forschung mitzuspielen.

Bedeutung des Standorts und von Core Facility Policy. Interessant ist, dass die MFPL sich eher den Forschungsinstituten am Standort Vienna Bio Center verbunden fühlen als ihren Schwesterinstituten an der Universität bzw. der MU Wien, deren gemeinsame Tochter die MFPL ja sind. Die Gründe dafür sind recht praktischer Natur: Man kooperiert entweder über den Gang oder mit dem Rest der Welt. Als Wiener nach Innsbruck zu reisen ist ebenso aufwändig wie etwa nach Heidelberg oder nach Cambridge (den beiden Standorten des *European Molecular Biology Laboratory (EMBL)*, um beim Thema zu bleiben). In diesen Kooperationen spielen sogenannte *Core Facilities* eine bedeutende Rolle. Auch hier kann etwa das *Vienna Bio Center* mit seinem *Campus Science Support Facilities (CSF)*¹⁰ eine Infrastruktur und eine *Core Facility Policy* vorweisen, die von europäischem Rang ist. Diese hohe Qualität kommt dadurch zustande, dass Forschungsagenden, (Rekrutierung von) Schlüsselpersonen und Investitionen in Forschungsinfrastruktur sowie deren Betrieb fein aufeinander abgestimmt sind und systematisch bewirtschaftet werden. Dies hat zur Folge, dass entgegen einer landläufigen Auffassung von Forschungsinfrastruktur, die Betreiber selbst die wichtigsten Nutzer sind – und umgekehrt.

Forschungsmanagement als the difference which makes the difference. Der Unterschied liegt weniger in den Prinzipien selbst als vielmehr darin, dass es wenige sind und dass diese wenigen ernst genommen werden: Personalpolitik, Kooperationspolitik, *Core Facility Policy*. Die Personalpolitik zeichnet sich vor allem dadurch aus, dass großer Aufwand für Rekrutierung und Karriereentwicklung getrieben wird, dass insbesondere Karriereabschnitt und Vertragsdauer systematisch aufeinander abgestimmt sind und dass junge Leute nicht erst dann angeworben werden, "wenn man ein Projekt gewonnen hat", sondern umgekehrt: Man sucht die Leute primär deswegen aus, weil man mit ihnen arbeiten will und bemüht sich in der Folge um die entsprechende Finanzierung. In der Kooperationspolitik steht nicht nur Akquisition und Durchführung gemeinsamer Projekte zur Diskussion, sondern auch Personalaustausch sowie Anschaffung und Nutzung von Infrastrukturen. Nicht zuletzt genießt *Core Facility Policy* eine besondere Aufmerksamkeit,

7 <http://www.gen-au.at/index.jsp?lang=de>

8 ELSA = ethical, legal, social aspects

9 Katharina Warta, Fritz Ohler, Cornelia Konlechner, Alfred Radauer, Barbara Good, Christien Enzig, Tobias Dudenbostel (2014), Evaluierung des österreichischen Genomforschungsprogramms GEN-AU unter Einbeziehung der Entwicklungen der Life Sciences Forschungslandschaft in Österreich, Technopolis Group Austria, 2104, abzurufen unter http://www.fteval.at/de/evaluation_studies/programmes/2014/1585/

10 <http://www.csf.ac.at/csf-facts/about-csf/>

vor allem durch die feine Abstimmung zwischen Forschungsagenden, Investitionen und der Mobilität von Personal.

Diese jungen Institute, das *Institute of Science and Technology Austria* sowie das *Institute of Quantum Optics and Quantum Informatics (IQO-QI)* sind hier noch einzuschließen, sie bilden einen Kern in der österreichischen Forschungslandschaft, der Anlass gibt, über zwei Phänomene nachzudenken. Das eine ist eine Quelle der Freude, das andere ist Anlass zur Sorge: Es ist offenbar möglich, erfolgreich Greenfield Investments in der Forschung zu tätigen, d.h. eine Investitionsentscheidung trägt nach 5-10 Jahren reiche Früchte. Sorge bereiten dürfte indes die sich abzeichnende institutional divide als Folge des beobachteten Matthäus-Effekts.

DAS ENTDECKEN DER BEDEUTUNG VON MANAGEMENT UND GOVERNANCE ALS SERENDIPITÄTSEFFEKT IN DER EVALUIERUNG

Die Entdeckung von Governance und Management als Schlüsselfaktoren ist ein anderes Beispiel für das Auftauchen des Serendipitätseffekts in der Evaluierung. In den *Terms of Reference* kamen die beiden Themen nicht zentral vor, und es wurde ihnen auch nicht viel Aufmerksamkeit in der Konzeption des Evaluierungsvorhabens geschenkt. Sie kamen vielmehr erst während der Interviews und vor allem während der Vor-Ort-Besuche nach und nach zum Vorschein, wo etwa die Institutsleiter/innen immer wieder ihre Managementprinzipien darlegten und durch Beispiele illustrierten, wo diese zu welchen Ergebnissen geführt hatten. Oder wo an einer anderen Stelle vom sukzessiven Zerfall eines größeren und sehr erfolgreichen Instituts berichtet wurde, weil die zwei zuständigen Dekane sich nicht zuständig fühlten und fünf Jahre bis zu einer Entscheidung durchs Land zogen, die am Ende halbherzig war. Schließlich zeigte sich, dass sich die erfolgreicheren Institute viel Zeit für unsere Evaluierung nahmen. Sie taten dies, weil bei ihnen selbst Evaluierung hoch angeschrieben ist, insofern dort alle ein, zwei Jahre für mehrere Tage, ja für eine ganze Woche der Betrieb still steht, weil eine intern angestoßene Evaluierung durch externe Gutachter/innen durchgeführt wird. Diese wird als etablierte Routine zur Sicherung von Qualität und Relevanz angesehen und steht weithin außer Streit.

Die maßgeblichen Erkenntnisse aus diesen Ausflügen in das Management und die Governance gehen weit über den Anlass der GEN-AU-Evaluierung hinaus. Sie waren weder bestellt noch drängten sie sich allzu sehr auf. Erst eine Reihe von massiven Problemen, ja Versagen an manchen Stellen und die Faszination über gelingende Forschung an anderen Stellen hat die Aufmerksamkeit der Evaluator/inn/en auf Fragen des Managements und der Governance gerichtet. Dazu kam auch ein biographischer Faktor hinzu: Der Autor dieses Beitrags hatte in den Jahren zuvor mehr als 50 Forschungseinrichtungen im akademischen wie industriellen Umfeld evaluiert, Leistungsvereinbarungen verhandelt oder Suchkomitees geleitet und war auf solche Fragen vorbereitet.

WAS KANN AUS DER ENTDECKUNG DES SERENDIPITÄTSEFFEKTS IN EVALUIERUNGEN FÜR DIE SPEZIFIKATION FÜR EVALUIERUNGEN GELERNT WERDEN?

1. Ressourcen für Serendipitätsentdeckungen bereitstellen. Die Auftraggeber von Evaluierungen sollten einen Teil ihrer Ressourcen für Serendipitätsentdeckungen widmen, die Evaluatoren davon kreativen Gebrauch machen. Wohin sollte der "vorbereitete Geist" aber seine Aufmerksamkeit bevorzugt richten?
2. Die Erkundung des Umfelds lohnt in jedem Fall, also das Erkunden verwandter bzw. komplementärer Programme bzw. Politikfelder einerseits, und des Kontexts der Forschung andererseits. Das ist schwierig genug, weil hier Reviere anderer Programme bzw. Agenturen und Institutionen betreten werden. Auch kennen Projektleiter/innen das weitere Umfeld in ihrem Unternehmen oder Institut oft nicht. Wegen des potenziell hohen Ertrags lohnen sich aber die entsprechenden Anstrengungen. Hier ist also Hartnäckigkeit geboten, weil Kontextualisierung hilft, die Evaluierungsergebnisse zu "erden".

Kontextualisierung kann auch ein Stück weit die so genannte *project fallacy* überwinden¹¹. Diese handelt vom irrigen Glauben, dass Antragssteller Projekte durchführen und Förderer Projekte fördern. Vielmehr werden durch geförderte Projekte früher getroffene Entscheidungen zusätzlich unterstützt und diese früher getroffenen Entscheidungen und die Arbeit daran werden später beendet als das geförderte Projekt. Die Folge ist, dass auch andere Finanzierungen eine bisweilen wichtige Rolle spielen können und dass das Fördergeld auch für andere Zwecke als die Durchführung des vereinbarten Projekts verwendet wird. Fördergelder werden also gewöhnlich "mitgenommen". Man kann diese zunächst enttäuschende Erkenntnis ein Stück weit aufhellen, wenn man den Kontext besser versteht und die Rolle des geförderten Projekts darin.

3. Fragen nach Management und Governance sind immer ergiebig. Für den "vorbereiteten Geist" ist die Formel "*management / governance is the difference which makes the difference*" die maßgebliche Suchstrategie. In dem hier diskutierten organisatorischen Kontext – Universitäten und öffentlichen Forschungseinrichtungen – ist vor allem das Mülleimer-Modell der Organisation eine hilfreiche Heuristik¹². Es beschreibt Organisa-

11 IWT Studies, Contributions to a Six Countries Programme Conference, February 28 - March 1, 2002, Brussels, http://www.iwt.be/sites/default/files/english/files/iwt_studie40.pdf

12 Michael D. Cohen, James G. March & Johan P. Olsen: A Garbage Can Model of Organizational Choice. In: *Administrative Science Quarterly*, Band 17, 1972, S. 1–25

tionen, die einen hohen Grad an lokaler, dezentraler Autonomie aufweisen, die unklare bzw. vage Ziele und Präferenzen haben und gleichzeitig mangelhaft ausgeprägte Verfahren und Regelungen, um zu Entscheidungen zu gelangen, und an denen nicht zuletzt das herrscht, was – im Original und schlecht übersetzbar – *fluid participation* heißt. Die Lösungsansätze, soweit es sie gibt, sind *Agenda Setting* als Voraussetzung zur Durchsetzung von Zielen, Führungspersönlichkeiten und manchmal auch Gremien¹³, die klare Ideen zum richtigen Zeitpunkt haben, nicht zuletzt aber auch operative Mitarbeiter/innen, die diese Ideen in der Praxis umsetzen.

4. Entdecken von *positive deviants*. Das, was in der Statistik gewöhnlich als Ausreißer wahrgenommen wird, ist in der Evaluierung oft eine so genannte *positive deviance*¹⁴, also ein Beispiel einer erfolgreichen Lösung in einem ansonsten widrigen Umfeld, mit dem Potenzial zur Verallgemeinerung. Es zwingt geradezu, in Kategorien von Kontexten und Management zu denken und in der Folge Lösungen zu entdecken, die ein Potenzial zur Verallgemeinerung haben. Die meisten statistischen Verfahren und der verführerische Fragebogen verdecken hier mehr als sie erhellen.

AUTOR

FRITZ OHLER

Technopolis Group Austria

E: fritz.ohler@technopolis-group.com

13 "Es geht immer um Gremien, aber die müssen auch einen Vor- und Nachnamen, eine Telefonnummer und eine Mail-Adresse haben, sonst kann man nichts ausrichten!" Ludwig Kronthaler, Generalsekretär der Max Planck Gesellschaft

14 "Positive deviance is an approach to behavioral and social change based on the observation that in any community, there are people whose uncommon but successful behaviors or strategies enable them to find better solutions to a problem than their peers, despite facing similar challenges and having no extra resources or knowledge than their peers." http://en.wikipedia.org/wiki/Positive_deviance

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

PETER KAUFMANN, IRIS FISCHL, SONJA SHEIKH

ABSTRACT

Begleitende Evaluation wie auch Begleitforschung werden als reflexiver Beitrag zur Steuerung im Programmvollzug verstanden, und erfolgen per Definition parallel zur Programmumsetzung. Gemäß den fteval Standards (2012) weisen sie einen eher formativen Charakter auf und fokussieren auf Lerneffekte für die beteiligten Akteure. Die zunehmende empirische Basis zeigt auf, dass diese Studien über die traditionelle Dreiteilung von ex-ante/interim/ex-post Evaluation insofern hinausgehen, als Elemente der Konzept-, Design-, Prozess- und Wirkungsanalysen im Zuge einer längerfristigen Begleitung von Programmen und Projekten je nach Anforderung flexibel integriert werden. Hierbei werden EvaluatorInnen mitunter (insbesondere in Deutschland) auch in die Implementationsunterstützung miteinbezogen. Neben der Lernfunktion zur unmittelbaren, schrittweisen Optimierung der Programmumsetzung wird auch ein hohes Potenzial für eine „prozessorientierte Wirkungsforschung“ gesehen.

Während diese Form der Programmbegleitung in der österreichischen FTI-Politik bislang noch relativ wenig vertreten ist (wenngleich mit steigender Tendenz), gehört sie in Deutschland bei neuen und komplexen Initiativen und Programmen inzwischen zum Standard.

Darüber hinaus schlagen wir eine eindeutige Begriffsbestimmung vor, indem sich Begleitforschung von begleitender Evaluation insofern unterscheidet, als letzteres umfassende Werturteile über das Evaluandum beinhaltet.

KEYWORDS: BEGLEITENDE EVALUATION, BEGLEITFORSCHUNG, ACCOMPANYING EVALUATION

In diesem Artikel wird die Frage nach dem Potenzial und den Herausforderungen von begleitender Evaluation bzw. Begleitforschung nachgegangen, indem das Thema auf Basis einiger konzeptioneller Überlegungen und eines Screenings bereits durchgeführter begleitender Evaluationen und Begleitforschungen in Deutschland und Österreich aufgearbeitet wird. Die Erkenntnisse wurden mit Interviews mit EvaluatorInnen ergänzt, um interessante Evaluationsdesigns und potenzielle Rollenkonflikte besser erfassen und darstellen zu können.

VERSUCH EINER BEGRIFFSDEFINITION

Eine Begriffsdefinition für die Begleitforschung sowie die begleitende Evaluation ist insofern nicht eindeutig, als die konzeptionelle Diskussion manchmal auf Unterschiede verweist, die Empirie jedoch zeigt, dass die Begriffe oft synonym verwendet werden. Darüber hinaus zeigen die Anwendungen der letzten Jahre, dass unterschiedliche Kontexte und Interessen ein recht hohes Ausmaß an variablen Inhalten und Rollenverteilungen nach sich ziehen. Es ist jedoch sowohl der begleitenden Evaluation als auch der Begleitforschung gemein, dass sie die Umsetzung eines Programms über einen längeren Zeitraum, bei sehr komplexen Programmen vereinzelt sogar während der gesamten Programmperiode begleiten. Dabei werden zur Entscheidungsunterstützung Informationen mit Hilfe des Methodensets der empirischen Sozialforschung aufbereitet und zum Zweck der Programm- und Politikentwicklung den beteiligten Akteuren unmittelbar zurückgespiegelt. Der Erfolg wird sowohl bei der begleitenden Evaluation als auch der Begleitforschung an der praktischen Nutzbarkeit der aufbereiteten Inhalte festgemacht.

UNTERSCHIEDEN SICH BEGLEITFORSCHUNG UND BEGLEITENDE EVALUATION?

Zu aller erst muss die Begleitforschung zur Umsetzung von Förderprogrammen von Begleitforschung für andere Zwecke abgegrenzt werden, wie z.B. dem identen Begriff zur Entwicklung konkreter Technologien. Bei letzterem werden, in der Technikfolgenabschätzung, hauptsächlich mögliche positive und negative Effekte einer Technologie auf die Gesellschaft und Umwelt untersucht.

Eine weitere Abgrenzung ist hinsichtlich jener Begleitforschung nötig, die separat beauftragte, programmbegleitende Forschungsprojekte im Rahmen von Programmumsetzungen darstellen, und z.B. bei der österreichischen Forschungsförderungsgesellschaft (FFG) als begleitende Maßnahmen bzw. Studien geführt werden. Diese Studien verfolgen meist das singuläre Ziel, begleitend zur Programmumsetzung die Marktentwicklung zu beobachten oder spezifische Bedarfe der Zielgruppe zu analysieren.

In der angewandten Sozialforschung wird manchmal zwischen Begleitforschung (alternativ: wissenschaftliche Begleitung) und beglei-

tender Evaluation¹ unterschieden. Während die *Begleitforschung* als eine alleinige Forschungsaufgabe unter dem Postulat der Wertfreiheit angesehen wird, hat die *begleitende Evaluation* dieselben methodischen Anforderungen zu erfüllen, aber zusätzlich eine bewertende Dimension (Kromrey 2008). Da sich die Beurteilungskriterien für die Zielerreichungsanalyse einer Evaluation aus den operationalisierten Programmzielen ableiten, ist hier noch keine explizite Wertung enthalten. Die Bewertung erfolgt dann zur Beantwortung spezifischer Evaluationsfragen, und darauf aufbauend bei der Entwicklung von Handlungsempfehlungen.² Im gleichen Atemzug wird jedoch auch eingestanden, dass die Begleitforschung in der Praxis als eine Form der angewandten Forschung einzelne Aspekte einer Wertung aufweisen kann, da die Wertfreiheit in der angewandten (Auftrags-)Forschung manchmal nur schwer durchgängig aufrechtzuerhalten ist.

Es bleibt jedoch unserer Ansicht nach der Aspekt, dass die begleitende Evaluation, im Gegensatz zur rein wissenschaftlichen Begleitung, eine umfassende Erfolgskontrolle inklusive Bewertung darstellen sollte. Die (begleitende) Evaluation weist einen vorgegebenen, umfassenden Rahmen für die schlussendliche Bewertung der zu erreichenden Ziele auf, wie dies z.B. anhand diverser Wirkungsmodelle in Evaluationen veranschaulicht wird. Die Informationsaufbereitung sowie die Bewertung werden dabei den EvaluatorenInnen in Personalunion überlassen.

In der Empirie der letzten Jahre wurden nach obiger Definition, trotz anderslautender Titel, meist nur begleitende Evaluationen in der FTI-Politik vorgefunden. Aufgrund dessen wird im restlichen Text nur mehr fallweise zwischen den beiden Varianten unterschieden.

ANSPRÜCHE VON BEGLEITENDEN EVALUATIONEN

Die begleitende Evaluation wurde im letzten Jahrzehnt aufgrund des Bedürfnisses von Programmträgern (insbesondere in Deutschland und auch auf der Ebene der Europäischen Kommission - Strukturfonds - in Form von ständigen (ongoing) Evaluationen) verstärkt umgesetzt, da Interimsevaluationen zumindest auf EU Ebene entweder zu spät zur Entscheidungsunterstützung durchgeführt wurden oder relativ wenig Einfluss auf die weitere Programmumsetzung zeigten. Von einer begleitenden Evaluation versprach man sich (i) einen stärkeren Fokus auf die strategische Ausrichtung des Evaluandums, (ii) ein Lernen im System mit unmittelbaren Feedbacks zur Verbesserung der Effektivität und Effizienz der Umsetzung und damit (iii) einen flexibleren Zugang, der mehr auf die wahren Bedürfnisse der Administration bzw. der Programmumsetzer eingeht. Da begleitende Evaluationen ihrer Natur gemäß über einen längeren Zeitraum implementiert werden, können die darin umgesetzten Teilstudien bzw. Arbeitspakete unterschiedliche Foci legen, und damit eine offensivere und an den Bedarfen des „Lebenszyklus“ einer Intervention ausgerichtete Rolle einnehmen.

So wird z.B. in den EU Strukturfonds den Nationalbehörden bezüglich einer Zwischen- oder begleitenden (hier: ‚ongoing‘ bzw. laufende)

Evaluation freie Hand gelassen, damit den internen Bedürfnissen der Programme besser entsprochen werden kann: „The advantage of ongoing evaluation is that it allows for effective collaboration between the evaluator and programme managers, which in turn favours a better appropriation of conclusions and recommendations“ (Evalsed, 2013: 107).

Wie dieser Zugang die formulierten Ansprüche erfüllt, wird in der Folge anhand eines Screenings von End- und Zwischenberichten zu begleitenden Evaluationen aufgezeigt.

SCHWERPUNKTE VON BEGLEITENDEN EVALUATIONEN VON FTI-POLITIKEN IN DEUTSCHLAND UND ÖSTERREICH³

Die Begriffe begleitende Evaluation und Begleitforschung/wissenschaftliche Begleitung werden in Deutschland synonym verwendet, obwohl entsprechende Analysen gemäß der obigen Definition meist den Charakter von begleitenden Evaluationen aufweisen. In Österreich scheint ausschließlich der Terminus begleitende Evaluierung auf.

In Deutschland ist der Einsatz von begleitenden Evaluationen bei neuen Initiativen bzw. Pilotförderprogrammen insbesondere auf der Bundesebene zum Standard geworden. Dies reicht von der Begleitung der Umsetzung der Pilot-Förderprogramme aufgrund der Hightech-Strategie (über jeweils 3-7 Jahre: KMU-innovativ, Spitzencluster-Wettbewerb, Validierung des Innovationspotenzials wissenschaftlicher Forschung – VIP), über diverse Länderprogramme, bis zur langfristigen Begleitungen von sehr komplexen Programmen wie z.B. die bislang 15-jährige Begleitung des EXIST Programms. Fallweise werden mit begleitenden Evaluationen auch bereits lang bestehende, relativ komplexe und mit einem gewissen Reformbedarf behaftete Programme adressiert (z.B. die erweiterte Erfolgskontrolle der Industriellen Gemeinschaftsforschung).

In Österreich wurden noch verhältnismäßig wenige begleitende Evaluationen im FTI-Bereich durchgeführt; diese sind aber beinahe ausschließlich innerhalb der letzten 10 Jahre angesiedelt. Die Initiativen betreffen entweder neue Zielgruppen in einem komplexen Umfeld, oder neue Komponenten im Instrumentenportfolio mit anfänglichem Entwicklungsbedarf (z.B. die Begleitungen der Kreativwirtschaftsprogramme, der Laura Bassi Centres of Expertise, des Sicherheitsforschungsprogramm KIRAS oder die Begleitung der Initiative innovatives-oesterreich. at bzw. der ‚excellentia‘ Initiative). Weitere Programmbegleitungen laufen derzeit oder sind in Vorbereitung (z. B. K-Zentren, Smart Cities,

1 Insbesondere im englischsprachigen Raum und auf der Ebene der Europäischen Kommission wird noch zwischen ‚accompanying evaluation‘ und ‚ongoing evaluation‘ unterschieden. Während erstere die Übersetzung des Begriffs begleitende Evaluation darstellt, ist mit ‚ongoing evaluation‘ der Prozess einer institutionalisierten, dauerhaften Begleitung z.B. der EU Strukturprogramme angesprochen (EC 2007, Evalsed 2013). Inhaltlich unterscheiden sich die beiden Zugänge jedoch kaum.

2 Für den Anspruch von Bewertungen mit ‚intersubjektiver Geltung‘ existiert nach Kromrey (2008) noch keine überzeugende Lösung.

3 Die empirische Basis an begleitenden Evaluationen ist der Literaturliste zu entnehmen.

innovative Subformate innerhalb von Mobilität der Zukunft). Die bereits fertiggestellten Begleitungen waren jeweils auf 2-4 Jahre angelegt.

Die Konsortien in Deutschland bestanden nur in einem Fall aus einer Institution, manchmal aus zwei, häufiger jedoch aus drei oder vier Partnerinstitutionen. In Österreich waren bislang ein bis drei Partner in eine begleitende Evaluation involviert. Für die im Vergleich zu Deutschland tendenziell geringere Konsortiumsgröße dürfte vor allem die Form der Ausschreibungen und die für die Evaluation verfügbaren Budgets ausschlaggebend gewesen sein.

INHALTE

Die Inhalte der untersuchten begleitenden Evaluationen integrieren jene der herkömmlichen Formate der Design-, Prozess- und auch Wirkungsanalysen. Typischerweise, jedoch mit unterschiedlichen Schwerpunktsetzungen, umfassen diese (1) eine Positionierungsanalyse des Programms im vorwiegend nationalen Programmportfolio inklusive einer Analyse der Interventionslogik, (2) eine Implementationsunterstützung und -analyse von Prozessen (Zielgruppenerreichung, Vernetzung/Moderation, Beratung, Transfer von Erkenntnissen nach innen und außen, seltener auch das Projektauswahlverfahren) und von etablierten Strukturen (z.B. die Funktionsweise von Netzwerken), (3) eine Zielerreichungsanalyse, und schließlich bereits (4) eine Wirkungsanalyse. Diese Themen werden manchmal ergänzt durch das Herausarbeiten von Good Practice Beispielen für besonders gelungene Implementationen sowie von branchen- bzw. technologiefeldspezifischen Qualifikationsbedarfen. So zeigte auch das INNO-Appraisal Projekt, eine Metaanalyse von europäischen Evaluierungen im FTI-Bereich auf, dass begleitende Evaluierungen die volle Bandbreite an evaluativen Fragestellungen und Methoden abdecken, und in Summe eine Kombination von bewertenden, lernenden und verbessernden Ansätzen darstellten (Edler et al. 2010).

Inhaltliche ‚Neuheiten‘ sind im Rahmen der Evaluationen hingegen nur vereinzelt zu finden: z.B. eine Analyse der Auswirkungen der im Förderprogramm entwickelten Technologien auf Gesellschaft und Umwelt (aus der klassischen Technikfolgenabschätzung); eine Komponente, die aufgrund der verstärkten Missionsorientierung von Förderprogrammen Schule machen könnte. Die Analyse von sozialen Wirkungen von FTI-Politiken wird fast ausnahmslos vernachlässigt, deren Berücksichtigung wäre jedoch im Sinne einer stärkeren Missionsorientierung durchaus wünschenswert.

Ein hoher Mehrwert der begleitenden Evaluationen liegt darin, das Monitoring sowie die weitere Informationsaufbereitung spezifisch auf das Zielsystem des Programms auszurichten, und damit sowohl kurze Lernschleifen zur unmittelbaren Weiterentwicklung der Initiative als auch mittelfristig eine tiefere Wirkungsanalyse mit hohem Anspruch ermöglichen zu können. Damit wird die Interventionslogik des Programms (theory of change) bereits während der Umsetzung zu weiten Teilen einer empirischen Prüfung unterzogen.

ANALYSE DES PROGRAMMUMFELDS

Bei einem relativ hohen Anteil an Interims- und Endevaluationen war in der Vergangenheit zu beobachten, dass die Analyse der externen Einflussfaktoren (Markt, Technologiefelder, politische Rahmenbedingungen) entweder gar nicht oder relativ oberflächlich umgesetzt wurde. Das mag damit zu tun haben, dass sich diese Art der Evaluationen

nicht (mehr) als Konzeptanalysen verstehen und somit auch nicht die entsprechenden Evaluationsfragen gestellt werden. Wenn ein Einfluss von sich verändernden Umfeldbedingungen vermutet wird, beauftragt das Programmmanagement fallweise separate Studien während der Programmumsetzung.

Da der Erfolg von komplexeren Initiativen oder Programmen in einem dynamischen Umfeld wesentlich von externen Faktoren beeinflusst werden kann, wäre es die Aufgabe einer begleitenden Evaluation, diese Einflussfaktoren systematisch darzulegen, und aufgrund dessen eine rationale Grundlage für eine entsprechende Studienvergabe zu etablieren bzw. die relevanten Informationen bereits im Rahmen der Evaluation selbst aufzubereiten und die Initiative / das Programm damit in einen längerfristigen Entwicklungskontext zu stellen. Diese Vorgehensweise konnte in den Berichten nicht nachvollzogen werden. Es stellt sich jedoch die Frage, ob dies überhaupt in jedem Fall dokumentiert werden würde.

ZIELANALYSE

Ziele werden in Programmdokumenten oft vage formuliert, da sie häufig entweder aufgrund von politischen Kompromissen oder Risikoscheue zustande kommen. Eine Präzisierung der Programmziele und deren Beziehungen zueinander ist jedoch sowohl für die Programmumsetzung ein wichtiges Signal, als auch eine (wünschenswerte) Bedingung für eine spätere Evaluation. Der richtige Zeitpunkt für eine konkrete Zielsetzung wäre in der Programmdesignphase. Da dies oft nicht passiert, kann eine Begleitforschung eine Konkretisierung der Ziele im Zuge der Programmumsetzung herbeiführen indem die analytische Denkweise der EvaluatorInnen mit den Realitäten der Programm-Stakeholder in Übereinstimmung gebracht werden. Hierbei sind die Kriterien der Konsistenz der Ziele zueinander, deren Operationalisierbarkeit, sowie deren Praktikabilität anzuwenden.

Da eine Zielanalyse für jegliche evaluative Tätigkeit nötig ist, erfolgt sie auch standardmäßig in Evaluationen jeglicher Art, wenn sie nicht bereits im Rahmen einer zeitlich vorgelagerten Arbeit durchgeführt wurde. Je später dies im Programmverlauf passiert, umso schwieriger kann es werden die zur Überprüfung nötige Informationen in ausreichender Qualität bereitzustellen. Da die Programmbegleitung meist zu Beginn einer Programmumsetzung ansetzt, kann sie das Zielsystem mit dem dazugehörenden Monitoringsystem sicherstellen.

KONTINUIERLICHER EINFLUSS AUF DIE PROZESSGESTALTUNG

Im Lebenszyklus eines Förderprogramms ergeben sich unterschiedliche Unterstützungs- und Informationsbedarfe. Nach einer strategischen Analyse sind etwaige programmunterstützende Aktivitäten und Informationsaufbereitung hinsichtlich der Zielgruppenerreichung, Projektunterstützung, Vernetzung innerhalb und über das Programm hinaus nötig. Hierbei sind auch Wechselwirkungen zwischen Projekten und Programmen (und damit potenzielle Mitnahmen von easy wins durch die Vernetzung von komplementären Initiativen) zu suchen, indem die EvaluatorInnen, zusammen mit den Programmverantwortlichen, entsprechende Prozessinterventionen gestalten.

Zur Steuerung der Lern- und Transferprozesse ist es von zentraler Bedeutung, geeignete Formate für Vernetzung und Ergebnistransfer zu etablieren. Aufgrund der großen Bandbreite an Formaten zur strukturier-

ten Stakeholderinteraktion ist die Auswahl heutzutage sehr hoch. EvaluatorInnen können dabei nur zwei Strategien verfolgen: (1) entweder ein Skilling-up an gruppenspezifischer Methodenkompetenz, oder (2) eine Kooperation mit einem Prozessberater einzugehen. Hier wird derzeit noch Potenzial gesehen, um die Ausgestaltung von Prozessbegleitungen noch lehrreicher zu gestalten.

In den untersuchten Evaluationen wird eine aktive Gestaltung und Moderation von Prozessen standardmäßig umgesetzt, wenngleich sie in einigen Fällen von etwas untergeordneter Bedeutung erscheint. Insbesondere wenn es um regionalwirtschaftlich relevante Programme geht, werden auch Formate wie offene Workshops und Konferenzen durchgeführt. Je neuer und komplexer die Initiative, desto partizipativer ist der Charakter der Begleitung. Hier wird eine Stimulierung durch Erfahrungsaustausch unter den geförderten Vorhaben und darüber hinaus gesucht.

Mit dieser Mehrfachrolle für die EvaluatorInnen geht aber auch einher, dass der Prozess der Programmentwicklung nun informell z.T. auch auf den Schultern der EvaluatorInnen lastet, und sie damit gegenüber dem Programm ein gewisses „Ownership“ aufbauen. Manchmal sind die EvaluatorInnen auch in die Beratung der Jury für die Projektbewertung und die Beratung des Ministeriums in Fragen des Fördergegenstandes eingebunden. Dies schließt auch die Unterstützung bei der projektübergreifenden Öffentlichkeitsarbeit mit ein. Durch die Erstellung von Kurzexerten wird im Rahmen der Evaluation besonders relevanten Themen auf eine flexible Art nachgegangen.

Der Vorteil einer intensiven Involvierung in den Prozess liegt darin, dass die EvaluatorInnen speziell bei komplexen Programmen und Portfolios diese besser verstehen und damit qualitativ höherwertige Handlungsanleitungen zur operativen Verbesserung des Evaluationsobjekts aus ihrer Arbeit resultieren können; d.h. der Nutzen für die Umsetzung der Evaluationsergebnisse wird damit erhöht. Der Nachteil besteht darin, dass die Objektivität der EvaluatorInnen darunter leiden kann und die Gefahr der ‚Vereinnahmung‘ durch die AuftraggeberInnen besteht. Damit würde eine begleitende ‚Evaluation‘ auf ein simples Modell der Politikberatung reduziert, die es nicht sein soll.

Diese Situation lässt sich nicht völlig befriedigend auflösen, man kann sich jedoch an der Praktik der Supervision, mit der in der Gruppendynamik bzw. Psychologie häufig gearbeitet wird, orientieren. Auch die eigene Erfahrung der AutorInnen zeigt, dass ‚Sparring-Partner‘ zur internen Reflexion, aber auch zur fallweisen Intervention gegenüber Auftraggebern eine sinnvolle Rolle einnehmen können. Im Falle eines Mangels an einer/eines geeigneten internen Person/Gremiums könnten durchaus auch institutsexterne ‚SupervisorInnen‘ für eine begleitende Evaluation vorgesehen werden.

DIE ‚PROZESSORIENTIERTE WIRKUNGSFORSCHUNG‘

Die Frage nach den Wirkungen der Programme und Initiativen wird in praktisch jeder Ausschreibung für eine begleitende Evaluation formuliert. Da Wirkungen für ihre Entfaltung bekanntlich ihre Zeit benötigen, stellt dies für eine Programmbegleitung von neuen Initiativen naturgemäß eine besondere Herausforderung dar. Diese können einerseits die Strukturen (Wirkungsmodell, Monitoring, Baseline-Surveys) für spätere ex-post Evaluationen legen, die dann aufgrund dieser Vorbereitungsarbeit eine höhere Qualität aufweisen werden. Andererseits

haben Programmträger bei neuen und innovativen Programmen auch die Erwartung, die kurzfristigen, qualitativen Effekte einzufangen, um erste Hinweise für etwaige Korrekturen im Programmedesign sowie der Umsetzung abzuleiten. Hier bietet sich natürlich das gesamte Spektrum der qualitativen Methoden der Sozialforschung an, ev. angereichert mit Formaten der (Groß-)Gruppenmoderation, um die qualitativer Erfassung von ersten Wirkungen voranzutreiben.

Es gibt jedoch auch begleitende Evaluationen von bereits länger laufenden Programmen oder die Prozessbegleitung ist über einen langen Zeitraum geplant; für diese kann durchaus eine ‚prozessorientierte Wirkungsforschung‘ umgesetzt werden. Dies ist auch in dem Kontext zu sehen, dass die Verallgemeinerung von Wirkungsketten aus individuellen Projekten hin zu Entwicklungen in Branchen oder Regionen derzeit in Wirkungsanalysen noch wenig mit nachvollziehbaren Kausalitäten belegt ist (siehe auch Kulicke et al. 2012). Der längere Zeitraum einer Begleitforschung sowie damit verbundene Budgetmittel könnten eine gute Voraussetzung für die dafür nötigen vertiefenden Analysen darstellen.

Die Erfassung von ersten Wirkungen während der Programmumsetzung bedeutet natürlich, dass noch keine umfassenden Aussagen möglich sind, aber sie kann bereits wertvolle Informationen hinsichtlich der Input- und Verhaltensadditionallität bei der Zielgruppe liefern.

In einigen Wirkungsanalysen im Rahmen der untersuchten deutschen begleitenden Evaluationen wurden tiefergehende Analysen angefordert, als dies in vielen Interims- und auch manchen ex-post Evaluationen der Fall ist: Das bedeutet z.B. die Verfolgung der zeitlichen Entwicklung von Wirkungsmustern und die Etablierung von Kontrollgruppen, deren Entwicklung über die Laufzeit der Programme verfolgt wurde, und die ebenso ungeplante kurz-, mittel- bis langfristige Wirkungen abzuschätzen versuchte. Dies wurde z.B. auf der Ebene von Projekten (qualitativ), von Akteuren (meist ein Matching bzw. Difference-in-Difference Ansatz), oder auch auf Clusterebene (mittels Peers, Stand der Technologieentwicklung) durchgeführt.

In Österreich standen in den bislang wenigen begleitenden Evaluationen mit einer Ausnahme Methoden des qualitativen und deskriptiv-quantitativen Spektrums im Vordergrund. Diese fokussierten stark auf die Vernetzung der ProgrammteilnehmerInnen, und damit weniger auf das Festhalten erster Programmwirkungen.

In Deutschland wurde als Datengrundlage nicht nur auf Monitoringdaten und Archivmaterial zurückgegriffen, sondern auch auf wiederholte eigene Befragungen der Zielgruppen und auch Kontrollgruppen, bzw. auf Sekundärstatistiken wie Patentamt, Mannheimer Innovationspanel, etc. Derartige Vergleiche mittels ökonomischer Methoden können aufgrund von Primärerhebungen oder aufgrund von Sekundärstatistiken durchgeführt werden. Die Verwendung von Sekundärdaten ist grundsätzlich zu bevorzugen, stellt sich jedoch manchmal als nicht durchführbar heraus. Besonders in Österreich ist das aufgrund der eingeschränkten Verfügbarkeit von Daten der Statistik Austria ein besonderes Problem, das behoben werden sollte.

Es fällt weiters auf, dass eine kreative Kombination von Methoden, wie bereits im Jahr 2000 durch Georghiou und Roessner angeregt, nur

langsam zu greifen beginnt. Auch schwerpunktmäßig qualitative Methoden (z.B. ein Pattern Matching im Rahmen von Fallstudien mit der Bildung von Kontrollgruppen, etc.) wären für die begleitende Wirkungsforschung durchaus vorstellbar, aber es konnte in den hier aufgenommenen FTI-Evaluationen keine Umsetzung identifiziert werden.⁴

EIN ADAPTIVES EVALUATIONSDESIGN ALS VORAUSSETZUNG?

Aus der empirischen Basis der letzten Jahre lässt sich ableiten, dass die Anforderungen an eine begleitende Evaluation sehr heterogen sind, und sich in den unterschiedlichen Evaluationsdesigns auch widerspiegeln. Aus Sicht der EvaluatorInnen ist aber auch während der Umsetzung Flexibilität notwendig, wenn auf interne Bedarfe der Programmträger während der Umsetzung reagiert werden sollte. Hier kann ein strikt vorgegebener Arbeitsplan beschränkend wirken.

Insbesondere wenn eine Initiative neu ist und in einem dynamischen Umfeld agiert, kann es ratsam sein nur die zentralen Aspekte der Begleitung genau zu definieren, und damit eine gewisse Flexibilität im Arbeitsplan und im Budget zu belassen, wenn später z.B. nicht vorhersehbare Expertisen oder andere Unterstützungsarbeiten benötigt werden sollten, oder andere bereits definierte Teile der Evaluation aufgrund des Gelernens nicht mehr so relevant erscheinen.

Um den vielschichtigen Prozess einer Programmumsetzung auf breiter Ebene unterstützen zu können, ist es ratsam einen ebenso begleitenden Expertenbeirat zu installieren, der sowohl Programm- als auch methodische Kenntnisse in sich vereint. Seine wichtigste Funktion ist die Reflexion von Zwischenergebnissen der Evaluation und die Unterstützung bei der Definition der folgenden Lernschritte. Je komplexer ein Programm und dessen Umfeld, desto wichtiger erscheint uns diese Funktion, da Entscheidungen getroffen werden müssen, wie knappe Mittel möglichst effektiv eingesetzt werden sollen. Damit wird nicht nur die inhaltliche Kompetenz der öffentlichen Auftraggeber ergänzt, sondern auch die flexible Handhabung des Arbeitsplans erleichtert.

BEWUSSTER UMGANG MIT WERTURTEILEN

Wie bereits weiter oben festgehalten, ist im Unterschied zur reinen Begleitforschung im Rahmen einer begleitenden Evaluation immer ein umfassendes Werturteil über das Evaluationsobjekt enthalten. Die Programmziele und die davon ableitbaren Kriterien und Indikatoren stellen die, durch die Politik vorgegebene, normative Basis dar. Damit erhält die Überprüfung der Zielerfüllung einen beinahe wertneutralen, deskriptiven Charakter. Aber von einer Evaluation wird noch mehr gefordert, nämlich eine gesamthafte Bewertung der Intervention.

Insbesondere wenn die EvaluatorInnen sehr stark in das Design und/oder in die (operative) Umsetzung der Intervention im Rahmen einer Programmbegleitung eingebunden waren, sind sie für das im Rahmen einer Evaluation schlussendlich abzugebende „Werturteil“ über die Sinnhaftigkeit der Initiative kompromittiert. Kromrey (2008) kommt aufgrund dieser Tatsache zu dem Schluss, dass die Prozessunterstützung und die Bewertung personell/institutionell getrennt werden sollten. Hierzu wären unserer Ansicht nach zwei Varianten denkbar: Entweder wird für die Projektend- oder Ex-Post-Evaluation eine andere Institution beauftragt als für

die begleitenden Evaluation, oder es übernimmt ein Expertenrat (auch mit internationaler Beteiligung) diese Aufgabe und bewertet die während der Evaluation aufbereitete Information. Die in die Prozessbegleitung eingebundenen EvaluatorsInnen arbeiten damit der End-Bewertung zu.

Demgegenüber kann eine Programmbegleitung mit relativ wenig operativer Involvierung durch die EvaluatorsInnen durchaus auch von diesen bewertet werden.

FAZIT

Befund 1: In Deutschland gehören begleitende Evaluationen bei neuen und komplexen Programmen bereits zum Standard, da man sich damit erhofft, dass im Vergleich zu anderen Formaten mehr auf die wahren Bedürfnisse der Administration bzw. der Programmumsetzer eingegangen wird. Diese Evaluationen weisen oft Laufzeiten von 3-7 Jahren auf, bei sehr komplexen Programmen vereinzelt sogar über die gesamte Programmlaufzeit und enthalten z.T. auch Elemente der inhaltlichen Programmunterstützung. Darüber hinaus wird bereits verstärkt auf eine „prozessorientierte Wirkungsforschung“, auch mit analytisch quantitativen Methoden, abgezielt. In Österreich ist diese leichte Veränderung in der Evaluationskultur erst in Entwicklung. Die noch relativ wenigen begleitenden Evaluationen sind tendenziell etwas kürzer angesetzt und fokussierten bislang überwiegend auf die unmittelbare Lernfunktion für die Programmumsetzung, und damit auf qualitative Inhalte. Um die Potenziale dieses Instruments voll auszuschöpfen, sollte der Fokus der Evaluation sowohl auf die Effektivität der Umsetzung als auch auf die prozessorientierte Wirkungsforschung gelegt werden.

Befund 2: Es bestehen definitorische Ungereimtheiten aufgrund der flexiblen Anwendbarkeit von Evaluation als Instrument der Politikberatung. Dies betrifft den Aspekt wer die Werturteile hinsichtlich des Erfolgs einer Initiative fällt. Während begleitende Evaluation eine Informationsaufbereitung und Beratung mit einer expliziten gesamthafte Bewertung darstellt, „fehlt“ der Begleitforschung dieses letzte Attribut. Damit geht auch eine unscharfe Abgrenzung von Inhalten und Rollen einher, die teilweise auf diese mangelnde Definition des Konzepts „begleitende Evaluation“, und andererseits auf unterschiedliche Kontexte und Bedarfe im Lebenszyklus eines Förderprogramms zurückzuführen sind.

Befund 3: Das Design der Evaluation sollte bei neuen und komplexen Initiativen insoweit als emergent angesehen werden, als bei einer begleitenden Evaluation einer (komplexen) Initiative eine gewisse Flexibilität im Design vorteilhaft ist, um auf Gelerntes oder neue Umfeldbedingungen reagieren zu können.

Befund 4: Unterstützende Aktivitäten bei der Programmdurchführung werden fallweise unter einer begleitenden Evaluation subsumiert (z.B. Involvierung in Bewertungsgremien, Koordination von Beratungsgremien, Beratung, Teile der Öffentlichkeitsarbeit). Je stärker EvaluatorsInnen in das Programmdesign bzw. in die Unterstützung der Politikintervention involviert sind, umso mehr legt dies eine institutionelle/personelle Trennung von Analyse- und Bewertungsschritten nahe.

⁴ Es könnte sein, dass die verstärkte Ausrichtung auf missionsorientierte Programme in Zukunft einen innovativen Schritt auslöst, da mit den gesellschaftlichen Handlungsfeldern ein breiteres Zielsystem im Vergleich zu den vorwiegend wirtschaftlichen Zielsetzungen der Vergangenheit berücksichtigt werden müssen. Dies hat auch Auswirkungen auf die zu verwendenden Methoden.

Damit wären die Hauptaufgaben der begleitenden Evaluation die Analyse der und Beratung zur Maßnahmenumsetzung inklusive das Sammeln erster Indizien für Input- und Verhaltensaddionalität, und das Vorarbeiten für die Endevaluation.

Die Bewertung an sich würde entweder von anderen EvaluatorenInnen oder von einem legitimierten Gremium übernommen werden. In der FTI-Politik bieten sich nach den Kriterien der Überparteilichkeit/Unabhängigkeit/Glaubwürdigkeit sowohl inländische also auch ausländische ExpertInnen, mit und ohne beratende Unterstützung durch die EvaluatorenInnen, dafür an.

Der Begriff begleitende Evaluation würde somit für den Prozess verwendet, in dem entweder Unterstützung, Analyse und Bewertung in Personalunion von den EvaluatorenInnen durchgeführt wird (bei relativ geringer Kompromittierung), oder auch wenn diese personell getrennt werden.

Befund 5: Manche Stakeholder sind besorgt, dass begleitende Evaluationen einen potenziell negativen Einfluss auf die (ohnehin nicht hohe) Reflexivität des FTI-Systems nehmen könnten. Grundsätzlich scheint uns dies eine Abwägung des vermeintlichen Nutzens aus einer verbesserten Umsetzung (und einer potenziell fundierteren Wirkungsanalyse) von komplexen Programmen und dem etwaigen Nachteil aus einer Vereinnahmung der EvaluatorenInnen zu sein. Während ersteres fallweise ein hohes Potenzial aufweist, muss dieses Potenzial erst durch ein geeignetes Design der Evaluation und eine profunde Implementierung gehoben werden. Wie oben beschrieben, können bezüglich der Dilemmas institutionelle Kontrollmechanismen eingeschoben werden. Hier ist auch die Transparenz der Prozesse und Entscheidungen angesprochen. Beispielsweise könnte man für wichtige Programmbegleitungen auch andenden, periodische Informationsveranstaltungen im Rahmen der fteval Plattform durchzuführen, um interessierte ExpertInnen in eine Reflexion der Zwischenergebnisse einzubinden.

Zusammenfassend kann sowohl Begleitforschung als auch die begleitende Evaluation Funktionen einer Konzeptions-, Implementations- sowie ‚prozessorientierten Wirkungsanalyse‘ übernehmen. Dies beinhaltet die folgenden Aufgaben: Sie stellt konzeptionelle Überlegungen bezüglich des Zusammenhangs von Zielen und dem Design der Intervention an, überprüft die abgeleiteten Thesen empirisch im Längsschnitt, beobachtet den Umsetzungsprozess im Bezugssystem der involvierten und betroffenen Akteure und Adressaten (z.B. Zielgruppenerreichung) und wertet ihn systematisch aus, unterstützt das Monitoring und ist aufgrund des Längsschnittcharakters auch in der Lage kausale Wirkungsketten „prozessorientiert“ zu untersuchen. Darüber hinaus kann man im Rahmen einer begleitenden Studie dieser Art auch flexibel auf (politische) Anforderungen und neue Erkenntnisse reagieren, sofern die Auftraggeber dies zulassen. Wenn gewünscht, transferieren Begleitstudien die gewonnenen Erkenntnisse nach innen und außen, und stellen die Entwicklung des Untersuchungsgegenstandes in einen längerfristigen Entwicklungskontext.

Der Einsatz einer begleitenden Evaluation/Begleitforschung ist insbesondere bei neuen und komplexen Initiativen sinnvoll, d.h. Modellinitiativen oder -programmen, deren Passform von Interventionsmechanismen und Art der Unterstützung sowohl konzeptionell entwickelt

als auch empirisch untersucht werden sollen. Weitere sinnvolle Anwendungsgebiete stellen längerfristige Programme mit eher diffusen bzw. komplexen Wirkungszusammenhängen dar, und/oder Programme, die lange Wirkungszeiträume aufweisen. Methodisch können begleitende Evaluationen durchaus Elemente von klassischen Evaluationen aufweisen, diese aber häufig aufgrund des längeren Zeitraums und einer flexibleren Handhabung des Budgets noch genauer umsetzen. Anhand des Einsatzes von Längsschnittdesigns mit der Integration von qualitativen und quantitativen Methoden lassen sich kausale Wirkungsmechanismen noch genauer untersuchen und heben damit das Potential des Programmdesigns.

Da die weitere Entwicklung unserer Gesellschaft und damit auch des FTI-Systems als Teil dessen systemische Ansätze der Politik erfordert, sollten auch Evaluationen vermehrt systemische Analysen vornehmen. Diese bieten sich besonders für das Stadium von ex-ante Assessments an, sind aber auch während der Implementation des Förderportfolios relevant. Damit steht die Überlegung an, ob ex-ante und begleitende Evaluationen bei Programmen in einem komplexen Umfeld zusammengezogen werden sollten.

LITERATUR

Aschoff, B., Astor, M., Crass, D., Eckert, T., Heinrich, S., Licht, G., Rammer, C., Riesenberger, D., Rüffer, N., Strohmeyer, R., Tonoyan, V., Woywode, M. (2011), Systemevaluierung „KMU-innovativ“. Studie von ZEW, Prognos AG, Institut für Mittelstandsforschung (ifm) der Uni Mannheim im Auftrag des BMBF, Mannheim und Berlin

Daimer, S., Berghäuser, H., Braun, A., Korte, S., Kripp, K., Kulicke, M., Meyer, N., Rijkers-Defrasne, S. (2014), Begleitende Evaluierung der Fördermaßnahme „Validierung des Innovationspotenzials wissenschaftlicher Forschung – VIP“. Studie des Fraunhofer ISI und der VDI Technologiezentrum GmbH im Auftrag des BMBF, Karlsruhe

Dörflinger, A., Sheikh, S., Vogl, B. (2012), Externe Begleitung zur Initiative „evolve“ für die Jahre 2009-2011. Studie der KMU Forschung Austria im Auftrag des BMWFJ, Wien

Edler, J., Cunningham, P., Gök, A., Rigby, J., Amanatidou, E., Garefi, I., Buehrer, S., Daimer, S., Dinges, M., Berger, M., Schmidmayer, J., Guy, K. (2010), INNO-Appraisal - Perspectives on Evaluation and Monitoring, Brussels

European Commission (2007), Indicative guidelines on evaluation methods: evaluation during the programming period, working document No.5, Brussels

Evalsed (2013), The resource for the evaluation of social-economic development, Brussels

Fischl, I., Mayer, S., Schüle, R., Lemken, T., Madry, T., Meinel, U. (2009), Evaluierung des ‚ÖkoBusinessPlan Wien‘, 3. Programmperiode 2006 – 2009 - Synthesebericht. Studie der KMU Forschung Austria und Wuppertal Institut für Klima, Umwelt, Energie im Auftrag der Stadt Wien

fteval (2012), Standards der Evaluierung in der Forschungs- und Technologiepolitik, Plattform Forschungs- und Technologieevaluierung, Wien

Georghiou L., Roessner D. (2000), Evaluating technology programs: tools and methods, *Research Policy* Jg. 29, 657–678

Heckl, E., Dörflinger, A. (2014), Begleitende Evaluierung der Impulsaktion 'Laura Bassi Centres of Expertise'. Studie der KMU Forschung Austria im Auftrag des BMWFW, Wien

Hornschild, K., et. al. (2005), Das BMBF-Förderprogramm InnoRegio – Ergebnisse der Begleitforschung. Studie des DIW Berlin, EuroNorm, CEIS und artop im Auftrag des BMBF, Bonn

Kind, S., Kaufmann, P., Ehrenberg-Silies, S., Fischl, I., Hannicke, S., Hoppe, U. Wolf, L. (2014), Endbericht zur Erfolgskontrolle des Programms zur Förderung der Industriellen Gemeinschaftsforschung (IGF). Studie des iit – Institut für Innovation und Technik der VDI/VDE Innovation + Technik GmbH und der KMU Forschung Austria im Auftrag des BMWi, Berlin

Kromrey, H. (2001), Evaluation - ein vielschichtiges Konzept. Begriff und Methodik von Evaluierung und Evaluationsforschung. Empfehlungen für die Praxis. In: *SuB Sozialwissenschaften und Berufspraxis*, Jg. 24(2), 105-132

Kromrey, H. (2008), Begleitforschung und Evaluation - fast das Gleiche, und doch etwas Anderes! In: Michaela Glaser, Silke Schuster (Hg.): *Evaluation präventiver Praxis gegen Rechtsextremismus. Positionen, Konzepte und Erfahrungen*. Leipzig: DJI, 113-135

Kulicke, M. (2013), Ergebnisse und Wirkungen der Förderprogramme EXIST-Gründerstipendium und EXIST SEED. Studie des Fraunhofer ISI im Auftrag des BMWi, Karlsruhe

Kulicke, M., Wessels, J., Sheikh, S. (2012), Wirkungsmodelle als methodische Herausforderung in Evaluationen. Verhältnis von klassischen Evaluationen und anderen Zugängen zur Bewertung von Innovationspolitik. *Zeitschrift für Evaluation*, Jg. 11(2), 330-344

Patton, M.O. (2008), *Utilization-focused evaluation*, 4th edition. Thousand Oaks, CA: Sage

Pfirmsmann, O., Heinrich, St., Riesenberger, D., Dinges, M., Haller, R., Hofer, R., Streicher, G., Ecker, B. (2012 und 2014), Evaluierung des österreichischen Sicherheitsforschungsprogramms KIRAS. Prognos AG, Joanneum Research, und Institut für empirische Sozialforschung im Auftrag des BMVIT, Wien.

Radauer, A. Dörflinger, A. (2007), Zweite begleitende Evaluierung des „Impulsprogramm creativwirtschaft“ der aws. Studie der KMU Forschung Austria im Auftrag des BMWA, Wien.

Rothgang, M., Cantner, U., Dehio, J., Engel, D., Fertig, M., Graf, H., Hinzmann, S., Linshalm, E., Ploder, M., Scholz, A., Töpfer, S. (2014), Begleitende Evaluierung des Förderinstruments „Spitzencluster-Wettbewerb“ des BMBF, Studie des Rheinisch-Westfälischen Instituts für Wirtschaftsforschung (RWI), des Instituts für Sozialforschung und Ge-

sellschaftspolitik, der Friedrich-Schiller-Universität Jena und Joanneum Research, Essen.

Rothgang, M., et al. (2010), Erweiterte Erfolgskontrolle beim Programm zur Förderung der IGF im Zeitraum 2005 bis 2009. Studie des RWI und dem WSF Kerpen im Auftrag des BMWi, Berlin.

Sheikh, S., Wessels, J. (2010), Begleitende Evaluierungen – Methodische Herausforderungen und aktuelle Erfahrungen. *Zeitschrift für Evaluation*, Jg. 9 (2), 243-246

Steiner, R., Fischl, I., Warta, K., Jörg, L., Bacher, J. (2007), Begleitende Evaluierung der Kampagne innovatives-oesterreich.at. Studie der KMU Forschung Austria, Technopolis Austria, TNS Infratest Hamburg im Auftrag einer Plattform bestehend aus BMBWK, BMVIT, BMWA und dem Rat für Forschung und Technologieentwicklung, Wien

Warwick, K., Nolan, A. (2014), *Evaluation of Industrial Policy: Methodological Issues and Policy Lessons*. OECD Science, Technology and Industry Policy Papers, 16: OECD Publishing

Wroblewski, A., Leitner, A. (2011), *excellencia Evaluationsbericht*, Studie des Instituts für Höherer Studien im Auftrag des BMWF, Wien

AUTORINNEN

PETER KAUFMANN

KMU Forschung Austria / Austrian Institute for SME Research
E: p.kaufmann@kmuforschung.ac.at

IRIS FISCHL

KMU Forschung Austria / Austrian Institute for SME Research
E: i.fischl@kmuforschung.ac.at

SONJA SHEIKH

KMU Forschung Austria / Austrian Institute for SME Research
E: s.sheikh@kmuforschung.ac.at

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

KLAUS SCHUCH, ISABELLA WAGNER, ALEXANDER DEGESEGGGER

ABSTRACT

This article is based on the evaluation of the bilateral intergovernmental science and technology agreements (STA), which Austria concluded with several countries all over the world. The process, methodological approach and the findings of this evaluation study are described briefly. In this article, specific emphasis is put on the application of bibliometric case studies, which we consider a relevant complementary part of the entire methodological portfolio used during the evaluation of the STAs. The article describes our approach in scrutinising a few bilateral project-based networks supported by STA through tracing the co-publication history of these networks and by investigating if and how the co-publication output has been influenced by STA support. During interviews with the coordinators of STA networks we presented our ex-ante identified co-publication results in the format of network graphs. The visual material led interviewees to recall their work conducted under the funded STA networks. This facilitated and stimulated an intensified debate on the usage, value and shortcomings of the STA on project level. On programme level, this allowed us to combine relevant quantitative and qualitative data for more solid results.

INTRODUCTION AND METHODOLOGY

In the course of the last few decades Austria concluded several bilateral agreements to support science and technology (S&T) cooperation with partner countries all over the world (Buzeczki, 2004). They include 15 active intergovernmental S&T agreements with Albania, Bulgaria, China, Croatia, France, Hungary, India, Macedonia, Montenegro, Romania, the Russian Federation, Serbia, Slovakia, Slovenia, and Ukraine. In addition there are two non-active intergovernmental S&T agreements with Israel and Spain. S&T cooperation is also tackled under two official bilateral cultural agreements with the Czech Republic and Poland, while the one with Italy is non-active. This portfolio is complemented with three active Memoranda of Understanding (MoU) with Argentina, Korea and Vietnam (ERAWATCH 2013). Although the different forms of agreements ask for different administrative processes with different rules, rights and obligations, they all support the individual international short-term mobility of cooperating research teams. To reduce complexity within this article, we summarize these agreements under the term science and technology agreements (STA).

It is important to understand that the STA, which Austria concluded with partner countries do not support research projects per se, but only international exchange on top of nationally funded projects. These projects are, however, to some extent bilaterally aligned between the Austrian applicant and the partnering applicant from abroad. The aligned international project part has to be submitted for funding in parallel to the responsible authorities in Austria and abroad. Between 1997 and 2013, 2,529 project coordinators working in Austria (and thus projects) were supported. Since the team of the project coordinator usually involves around two more team members, who participate in some short-term mobility exchanges, several thousand researchers residing in Austria were supported. This indicates a comparatively high mobilisation effect given the small budgets involved (on average approximately €2,000 for around 4 trips per project and year). Between 2002 and 2012, €5.241m were spent under the STA framework in Austria.

In 2013 the former Austrian Federal Ministry of Science and Research commissioned the Centre for Social Innovation (ZSI) to evaluate the STAs along the following questions:

1. Positioning of the relevance of the STA vis-à-vis the (i) newly developed Austrian STI internationalisation strategy called "Beyond Europe" (Arbeitsgruppe 7a, 2013), (ii) the existing portfolio of internationalisation instruments and (iii) the communication published by the European Commission on RTI internationalisation (European Commission, 2012)
2. Assessing the efficacy of the STA along intrinsic R&D internationalisation objectives as well as along pertinent science diplomacy aspects
3. Scrutinising the efficiency of the different contractual types of STA (bilateral intergovernmental agreements versus other types of S&T agreements) compared to the involved administrative effort.

To meet these requests, the evaluation team deployed a set of regular standard methods including literature and document analysis, an online survey with Austria-based coordinators of funded bilateral STA networks for descriptive and analytical statistical analysis, as well as expert interviews with representatives of the science ministry, the ministry of foreign affairs (including embassies abroad) and the OeAD, the Austrian Agency

for International Cooperation in Education and Research, who is responsible for the administrative implementation of the STA. To complement this portfolio of methods additionally seven case studies of previously funded STA networks were conducted. The qualitative interviews with the Austria-based coordinators of these seven STA networks were based on an ex-ante investigation of the co-publication patterns involving the Austrian based coordinator and his/her team together with the foreign STA coordinator and his/her team. The basic idea of this approach was to verify whether the Austrian and foreign partners have co-published papers together and if the co-publication intensity has increased after approval of the bilateral STA funding.

The evaluation started in March 2013. The final report was adopted in September 2013, one month after the official publishing of the "Beyond Europe" strategy. 472 persons responded to the online survey, which represents a return rate of around 30%.

EXCERPT OF MAIN FINDINGS

The multi-method approach yielded a broad spectrum of findings (Schuch et al., 2013), of which just an excerpt is presented in this section to provide a basic understanding of the STA, their acceptance and use.

First of all, it can be stated that the prototypical recipient of an STA funding has reached an advanced career level and works at a university in the field of natural sciences. He¹ usually uses the STA funding for continuing an existing international partnership by creating added value through increased and qualitatively improved publication output. He is satisfied with the low bureaucratic effort caused by the STA funding and satisfied with the OeAD as administrative counterpart. He is criticising the ex-ante project proposal evaluation and the insufficient feedback culture in this regard.

The acceptance of STA funding is strongly determined by the funding design, which neither finances personnel costs nor overheads. Thus, research institutions without structural basic research funding are hardly encouraged to apply (e.g. non-university research organisations and universities of applied sciences). Most calls for proposals launched under STA are thematically open. Accordingly, the participants in STA networks shape the scientific and thematic orientation of STA, which can be characterised as basic research with a high bottom-up demand from natural and technical sciences. Only 11% of funded STA networks belong to social sciences and humanities.

Secondly, one can summarise that the portfolio of Austrian STA partner countries has grown historically by including "older" and "newer" EU member states, non-EU countries, which were associated to the 7th European Framework Programme for RTD, as well as "classical" international partner countries in Europe (Russia, Ukraine) and overseas (e.g. China, India). The range of partner countries has been growing continuously and is responding to the growing cooperation demand (and need) reflecting an accelerating globalisation in the field of S&T. During the last years, this globalisation trend is evidenced by an increasing outreach of the Austrian STA towards geographically more distant but

emerging countries such as Argentina (STA network funding since 2011), India (2011) and South Korea (2008). In parallel to this globalisation outreach, further investments were made within a wider neighbourhood policy through a number of new agreements with Macedonia (2011), Montenegro (2011) or Ukraine (2007).

Within the Austrian "Beyond Europe" strategy, potential (or active) partner countries for increased RTI cooperation have been clustered in three priority levels (Arbeitsgruppe 7a, 2013). Concluded STA agreements can be found at all three priority levels and even with non-prioritised countries. An increased usage of bilateral agreements is suggested by the strategy. As a low-threshold instrument with a strong mobilisation effect, we recommended a further use of STAs with countries subsumed under priority level two (if not yet realised), but especially with countries subsumed under priority level three as well as with countries whose potential for future collaborative RTI investments still needs to be explored ('door-opener function').

In contrast to this globalisation and wider neighbourhood tendency of new STA agreements, Austrian researchers participating in STA prefer the productive cooperation with "older" EU member states, not least because of a much better transition of the established cooperation into the European Framework Programme for RTD and other European programmes (Schuch et al., 2012). They also expect an even stronger internationalisation need and demand with Southeast Asian countries in the near future, which is not yet tackled sufficiently by STA agreements.

In view of the high mobilisation effects of the STA one can doubtlessly qualify them as cost-efficient. However, the exercised austerity also leads to subcritical team sizes, subcritical institutional effects and a singling of the supported networks due to the lack of a superordinate umbrella with accompanying framework activities capable to bridge to more substantial funding programmes.

The immediate effects of the STA can be summarised as follows:

- STA networks are catalytic for already established cooperation relations. They are less frequently used to initiate new cooperation relations, notwithstanding the indications from applicants that STA would also be helpful in this regard.
- The dominant benefit according to STA network members lies in the preparation and drafting of joint publications.
- Only a few other cooperation effects can be identified such as access to foreign research infrastructure or to data, which are not domestically available. Due to the small budgets, supported STA networks are hardly used to conduct field research, to conduct experiments, to establish contacts to commercial agents or to economically capitalise research results. One has to be fair enough to state that market penetration and capitalisation are not programmatic objectives of STA.

In terms of sustainability of the STA networks one can conclude that after termination of the STA funding

1 The majority of Austrian STA network coordinators are male.

- 89% of the survey respondents maintain contact with their foreign partners
- 61% continue to co-publish with their former foreign partners
- 50% conducted one or more follow-up projects

The results of the logistic regression model showed that those researchers have higher chances for a long-term cooperation after termination of the STA funding,

- who have already worked together with their project partner before their supported STA project
- who are very satisfied with the attainment of project goals during their previous STA project
- who have published an above average number of co-publications with their cooperation partners
- who cooperate with a partner from the “old” 15 EU member states

FINDINGS OF THE BIBLIOMETRIC CASE STUDIES

As already indicated in the introductory part of this paper, the range of effects of the STA networks was also scrutinised by a case study approach. Seven STA networks recommended by OeAD were further investigated through qualitative interviews to identify immediate scientific and socio-cultural outputs and wider outcomes of their cooperation. To substantiate the interviews and to prepare a discursive input for the conversations, we analysed the publication and co-publication output of the STA case studies under scrutiny before the interviews. The goal was to verify ex-ante how, if at all, STA funding materialised in tangible publication output.

This approach was documented and analysed along the following scheme:

- Investigation of the publications of each cooperating STA team (Austrian and foreign STA cooperation partners), taking into account the

- year of publication (before, during [plus 2 years] and after the bilateral STA project funding) and assessing the
- quality of publication (measured by number of citations)

First of all, this base set of bibliometric output was assessed through quantitative bibliometric indicators based on co-publications between the STA network partners recorded in the Scopus database. International co-authorship was operationalised as the definitive indicator. By employing name/institution-couples, all co-publications among the team members (as far as these individuals were known to us through the respective project applications and project reports) engaged in the specific STA networks were recorded. Beyond this, we also aimed at catching the leverage effect of the STA networks within Austria by tracing co-publications of the foreign partners with Austria-based team members other than the coordinator during or after the STA project (intra-organisational enlargement).

The results of this bibliometric approach towards STA networks were visualised by us in network graphs. These visualisations were presented to our interview partners (Austrian coordinators of STA networks) by using a Problem-Centred Interview (PCP) approach (Witzel and Reiter, 2012) to provide us with their qualitative assessments. In addition, the interviews were supported by a field manual. The following table 1 shows the scrutinised cases.

The researchers participating in the seven scrutinised STA networks published 53 joint publications. The appearance of these publications was not limited to the duration of the funded STA networks but included a timespan between 2002 and 2013, observing some years before the start and after the termination of the funded STA networks. We counted all publications which were mentioned in the final project reports, including those cases in which the Austrian based project coordinator was not among the authors, as well as co-publications authored by the Austrian and the foreign project coordinator recorded in the Scopus literature and citations database. However, we did not aim to identify all the publications of the team members through our Scopus database research. Between four and 13 joint publications could be identified per project; an average of 7.6 co-publications per investigated STA network.

For those publications indexed in Scopus, the frequency of citations in other publications recorded in Scopus were assessed (cut-off date May 2013). The number of citations received per publication ranged

STA-network number	Institution	Field of research	Partner country	Year
1	University of Natural Resources and Life Sciences, Vienna (BOKU)	Food science and technology	IT	2004
2	University of Vienna	Mineralogy & crystallography	RU	2004
3	University of Innsbruck	Ion physics	SK	2006
4	Zentralanstalt für Meteorologie und Geodynamik (ZAMG)	Meteorology	FR	2005
5	Technical University of Vienna	Material science and material technology	ES	2006
6	University of Graz	Geophysics, astrophysics. & meteorology	SK	2006
7	Medical University of Vienna	Biomedical research	PL	2006

Table 1: Cases for bibliometric profiles and target group interviews

between 0 and 55 and at average amounted to 9.2. For more recent publications, these numbers do not provide a robust indication about their long-term (scientific) impact. On the other hand, citations close to the publication date hint to the immediate relevance of research results for the respective scientific field ('immediacy'). Self-citations were not excluded because the citation of a joint output by one of the STA project partners could indicate the usage of jointly produced results within the national context, which we consider relevant in the context of our study.

Since the investigated STA projects belong to different thematic areas with different citation behaviour, we also list in Tab. 2 the average number of citations per co-publication with Austrian participation in the corresponding year and thematic area. Thus, we can compare if the STA related co-publications are above or below the theme-specific Austrian co-publication average. It became evident that most STA related co-publications are clearly producing an above-average impact, which indicates the excellence level of STA participants.

STA-network number	Year of publication	Scientific field (Science Metrix field or sub-field)	Times cited	Avg. Times cited per scientific field in the reference year (and number of Austrian co-publications in the reference year)
STA network number 1, with Italy 2004	2005 2008	Food science and technology	35 10	16.2 (29) [2006] 13.4 (47)
STA network number 2, with Russia 2004	2005 2006	Applied Physics	11 6 11	10.4 (263) [2006] 10.4 (263) [2006] 10.4 (263)
STA network number 3, with Slovak Republic 2006	2009 2010	Applied Physics	13 6	4.3 (241) 3.2 (276)
STA network number 4, with France 2005	2011	Meteorology & Atmospheric Sciences	13	2.7 (266)
STA network number 5, with Spain 2006	2011	Materials science	2	0.9 (173)
STA network number 6, with Slovak Republic 2006	2006 2008 2009 2010 2010	Geochemistry & Geophysics	47 55 35 13 9	7.9 (128) 8.6 (137) 3.5 (203) 2.3 (173) 2.3 (173)
STA network number 7, with Poland 2006	2006 2008 2009	Biomedical research / Biomedical engineering	8 14 8	24.2 (727) / 18.0 (40) 18.0 (854) / 6.1 (65) 12.7 (944) / 4.3 (98)

Table 2: Comparison of citation measures of selected publications with scientific field-related benchmarks

Another approximation for the level of excellence (including seniority as the h-index is higher for longer research careers) is the number of indexed publications per researcher and their Hirsch-index values (h-index). The Austrian STA network coordinators have between six and 187 publications indexed in Scopus (on average 113.3). For the foreign STA coordinating partners, the range is between 11 and 161 publications (on average 83.4). The h-index of the Austrian coordinators ranges between three and 33 (on average 18.2), while the range of the foreign coordinating partner is between eight and 22 (on average 12.7).

Finally it was also monitored how often the STA coordinators from Austria and abroad were citing each other during their scientific careers (operationalised in terms of all publications of the respective person indexed in Scopus). While the Austrian STA coordinator cites the foreign STA coordination partner at average 9.7 times, the reverse citation average is 14.6. Excluding cross-citations within a joint publication, the resulting average cross-citation rate reaches 4.3 as regards to the Austrian STA coordinator citing his or her foreign STA coordination partner versus 5.7 the other way round.

TWO CASE STUDIES FROM THE FIELD OF FOOD SCIENCE AND TECHNOLOGY AND GEOPHYSICS

Three out of 7 STA networks did not produce joint publications before the start of their networks. However, during and after the termination of the projects, the network members co-published and continued to co-publish in one or the other personal composition. The evolution of co-publishing activities over time is exemplified by the following two timeline graphs from the field of agro-food technology (shown in Fig. 1) and geophysics (shown in Fig. 2).

Within the first case study, the two network coordinators (one from Austria and one from Italy) had published two co-publications indexed in Scopus (2001 and 2003) already before the start of their joint STA network. They published three joint papers during the life span of their STA network (2004 and 2005). Two of those (in which also a team member of the Austrian coordinator was involved) were listed in the network's final report, but not indexed in Scopus. After the termination of the STA network, two more co-publications indexed in Scopus (2008 and 2012) were produced by the former network coordinators and a team member

Figure 1: STA case study 1 – IT 04/2004 BOKU, Vienna and Consiglio Nazionale della Ricerche, Milano, Italy

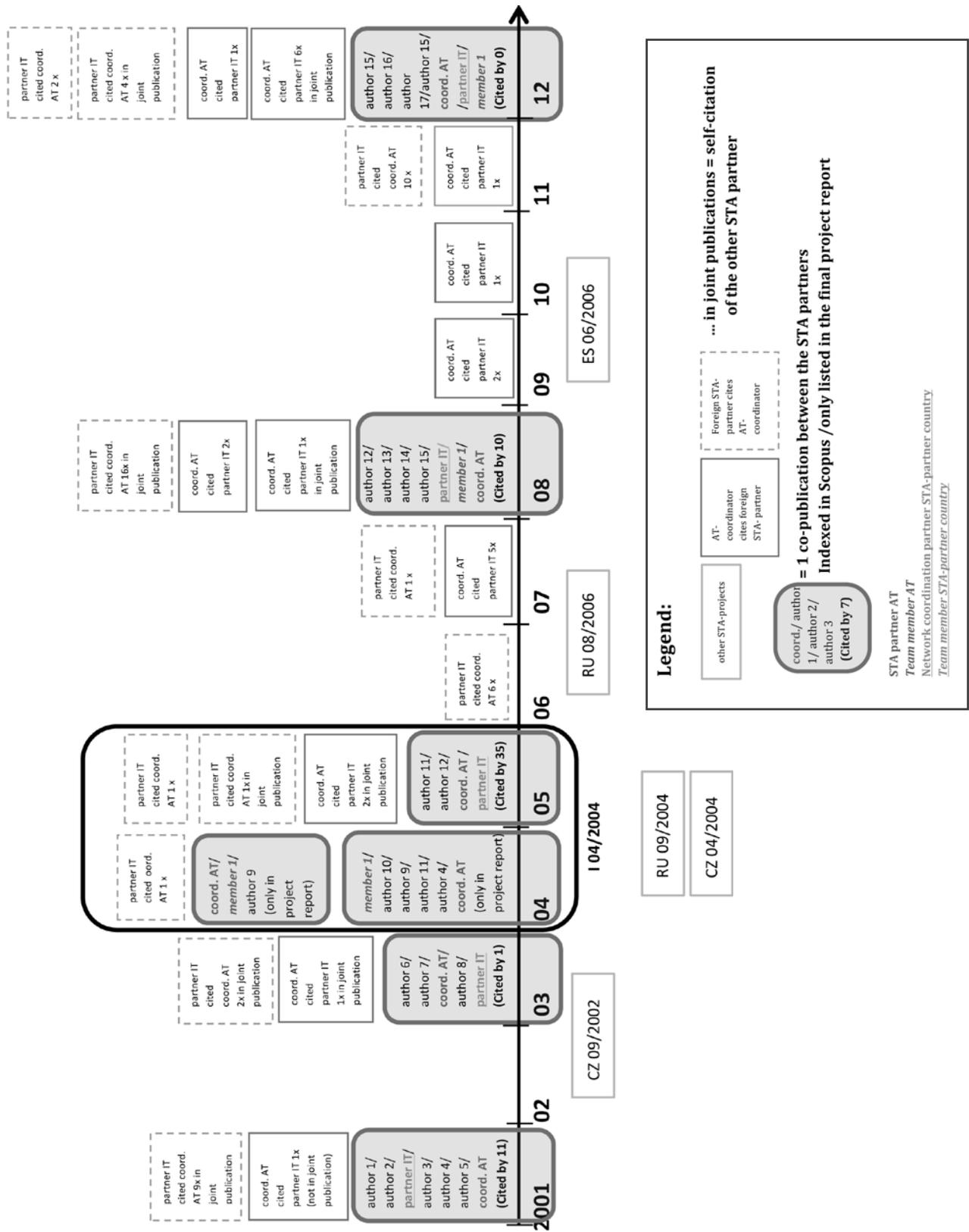
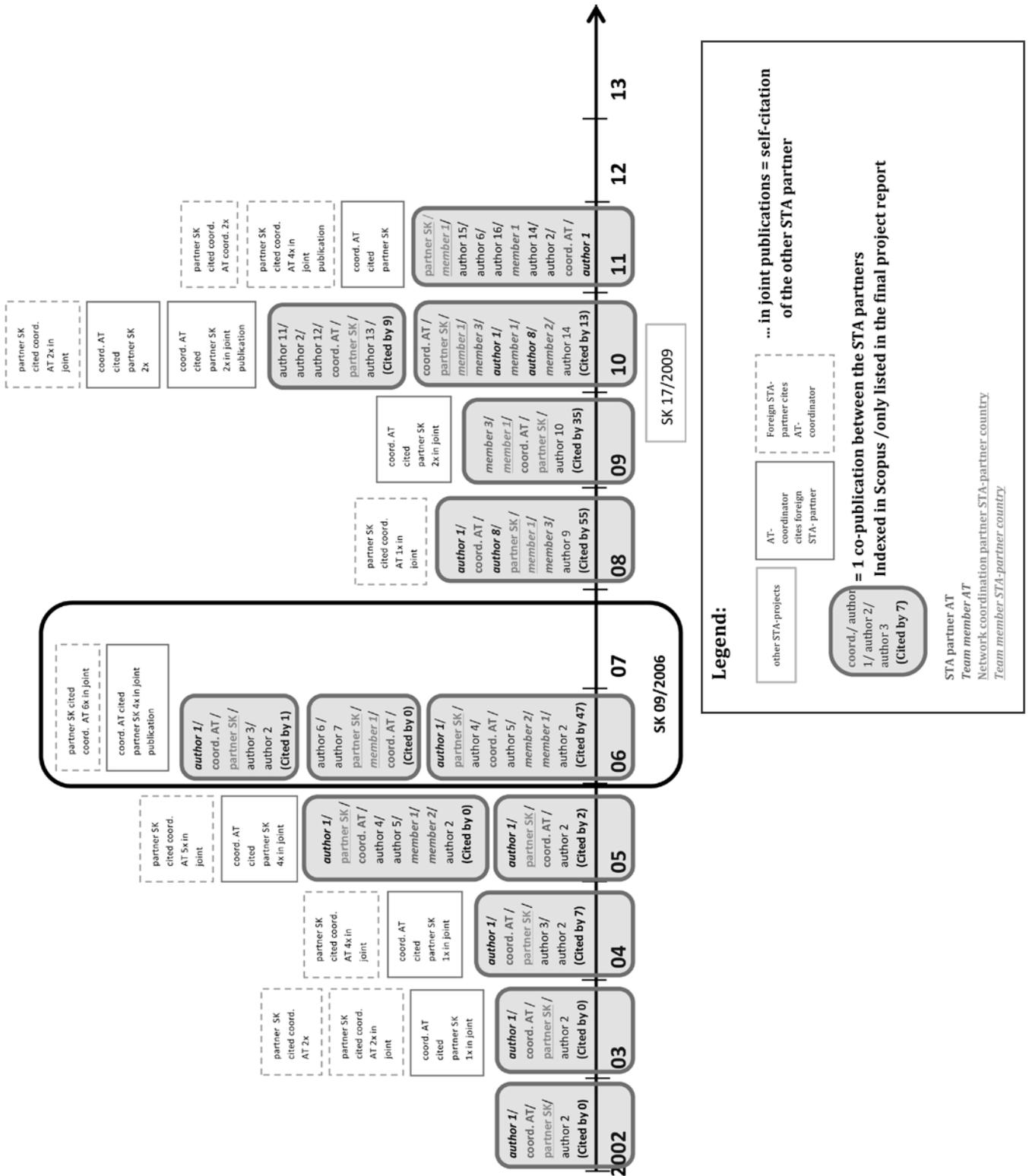


Figure 2: STA case study 2 - SK09/2006 University Graz and Slovak Academy of Sciences



of the Austrian coordinator. For the five co-publications indexed in total in Scopus, we counted how often they had been cited by other publications (see Fig. 1). It is not surprising that the most recent publication from 2012 had not yet been cited because of our retrieval cut-off date in 2013.

As shown in Fig. 1 the co-publication from 2008 was cited by 10 other publications. The one of 2005, which was produced during the life span of the joint STA network, was cited 35 times. The two co-publications, which were published before the joint STA network, were cited less frequently. The two network coordinators cite each other frequently. With a total of 76 reciprocal citations, of which 42 are made in joint publications, the two STA network coordinators are intensively referencing each other, which is an indication of a robust cooperation. Out of the five other STA network projects of the Austrian coordinator plotted in Fig. 1 two are with the Czech Republic and two with Russia, each with the same foreign cooperation partner.

The second case study is composed by an Austrian STA coordinator and his corresponding Slovak STA coordinator in the field of geophysics (solar physics). The co-publication history of this STA network is shown in Fig. 2. From the 7 cases scrutinised, this is the STA network with most co-publications ($n=13$). The two coordinators had already co-published five publications before their joint STA project. They produced three more co-publications during the STA project life span (2006 and 2007), and another five after the end of the STA project.

Most of the co-publications, which were produced during their first and until their second joint STA project, are cited more often compared to their co-publications published before 2006. While the later were cited between 0 and 7 times, most of the seven co-publications between 2006 and 2010 were cited more frequently; four more than 10 times and three even more than 30 times. The Austrian coordinator cited the Slovak coordination partner 18 times, out of which 14 citations were in joint publications, while the Slovak coordinators cited his Austrian counterpart 28 times (24 times in joint publications). Overall, they reference each other less frequently than the coordinators in the first case study. Moreover, most of their mutual citations are in their own co-publications, which therefore should be considered as self-citations.

The other five case studies revealed similar tendencies. The funded STA network underlying case study 3 had – according to our bibliometric observations – an initial impact and an integrative effect on the cooperating teams, but only in a short-term perspective. Case study 4 showed that the funded STA network was a link of a chain within an established, sustainable international partnership with an accelerating and integrative effect on the cooperating teams. The same could be observed for case study 5, however with slightly lower cooperation intensity. The STA network underlying case study 6 had initial impact on the cooperation, but an extension of co-publication output did not take place over time. In case study 7, finally, the STA network also had initial impact on the cooperation leading to a moderate, but sustainable co-publication output.

With these bibliometric findings, equipped with corresponding figures (as shown above) and a field manual, we approached these seven Austrian STA coordinators for a Problem-Centred Interview (Witzel and Reiter, 2012) to provide us with their qualitative assessment of our quantitative findings. We also asked them for a general reflection about the STA programme and some of its aspects. The majority of insights from the

qualitative interviews, which cannot be explicated in detail here due to the limited space of this journal, confirmed the findings, which we obtained by the survey. In addition, they enabled a more robust and complementary interpretation of survey data. To preserve the special character of the interviews, we summarised the interview findings in a dedicated chapter within the evaluation report (Schuch et al., 2013). The findings focussed on (i) the qualitative assessment of the co-publication outputs of the STA networks under scrutiny, (ii) the network development, (iii) the sustainability of cooperation, (iv) the programmatic relevance of the STA programme for the beneficiaries, (v) its instrumental contribution to internationalisation, and a few more aspects. They can be looked up in the final evaluation report accessible at the website of fteval (http://www.fteval.at/en/evaluation_studies/all/2013/1313/).

CONCLUSIONS

One can conclude that in the majority of observed case studies, the STA cooperation resulted in an evident increase of co-publication output between the participating network partners including some very visible publications with strong impact. As confirmed by the survey, the conceptualisation, preparation and drafting of (co-)publications is both a strong motive and an end of a STA project. This indicates that a successful STA project can be characterised by a quantitatively and qualitatively increasing joint publication output. Ideally, this co-authorship activity is also sustainable in the sense that the previous STA partners continue to produce joint publications even beyond the termination of the initial STA project. Within such beneficiary-confirmed interpretation it is not necessary that STA networks focus on new partnerships, but that STA networks can be likewise successful if they involve partners who would anyway cooperate, but whose cooperation is reinforced and condensed by the available mobility means provided by the STA programme. However, the interviewees often pointed to the fact that the production of publications should not be too simplistically allocated to the STA network funding only, but that the STA should rather be considered as just one input among others (see also Morris, 2003; Rigby 2011).

We can further conclude that bibliometric case studies are an inspiring complementary method for S&T programme evaluations, when the programme under scrutiny has a strong publication output orientation. The interview situation was enriched by an additional input (the network graphs), which created a volitional reference point for the interview. Because of this preparatory work, the interviewees also felt taken seriously. It was interesting to observe that some interviewees initially underestimated the productivity of their STA networks and were positively surprised when we confronted them with our findings.

It was also possible to identify a minor leverage effect of the STA networks in terms of co-publications of the foreign partners with Austrian based team members other than the coordinator (intra-organisational enlargement), although in general the STA cooperation seems to be strongly personalised and focused on the coordinators.

Only one of seven interview partners said that the funding provided by the STA programme is too small to produce a direct output of any kind. There was consensus that mobility in general increases international visibility, either of Austria as science location, of the own institution

or – most importantly – one’s own visibility and the visibility of the co-operating research groups through joint publications.

REFERENCES

Arbeitsgruppe 7a der FTI-Task Force „Internationalisierung und FTI-Außenpolitik“ (2013): Beyond Europe. Internationalisierung in Forschung, Technologie und Innovation. Strategiepapier und Aktionsplan. Wien: BMVIT und BMWF.

Buzeczki, C. (2004): Bericht zur Evaluierung der Mobilitätsförderung wissenschaftlicher Kooperationsprojekte im Rahmen der bilateralen Abkommen für wissenschaftlich-technische Zusammenarbeit des bm.bwk. Wien: bm.bwk.

ERAWATCH Network asbl. (2013): Country Report Austria of the Final Report: Overview of International Science, Technology and Innovation Cooperation between Member States and countries outside the EU and the development of a future monitoring mechanism (prepared by K. Schuch, Centre for Social Innovation). Brussels: ERAWATCH.

European Commission (2012): Communication on Enhancing and focusing EU international cooperation to research and innovation: A strategic approach; COM(2012) 497.

Morris, N. (2003): Academic Researchers as 'Agents' of science policy, *Science and Public Policy*, 30 (5), 359-370, 2003.

Reiter, Herwig and Witzel, Andreas (2012): The Programme of the PCI. In: *The Problem Centred Interview: Principles and Practice*. London: SAGE Publications Ltd., pp. 12-34.

Rigby, J. (2011): Systematic grant and funding body acknowledgment data for publications: New dimensions and new controversies for research policy and evaluation. *Research Evaluation* 20, pp. 365-375.

Schuch, K., Smoliner, S., Wagner, I., Degelsegger, A. und Dall, E. (2013): Evaluierung der Forschungskoperationen im Rahmen der bilateralen wissenschaftlich-technischen Abkommen und Vereinbarungen. Wien: Zentrum für Soziale Innovation. Studie im Auftrag des BMWF.

Schuch, K., Wagner, I and Dall, E. (2012): The potential of transfer of bilateral R&D projects towards the European Framework Programme for Research and Technological Development. In: Loudin, J. and Hochgerner, J. (2012): *Social and Cultural Dimensions of Innovation in Knowledge Societies*. Prague: Filosofia: pp. 169-197.

AUTHORS

KLAUS SCHUCH
Centre for Social Innovation, ZSI
E: schuch@zsi.at

ISABELLA WAGNER
Centre for Social Innovation, ZSI
E: wagner@zsi.at

ALEXANDER DEGELSEGGER
Centre for Social Innovation, ZSI
E: degelsegger@zsi.at

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

IRIS FISCHL, PETER KAUFMANN

ABSTRACT

Die Bandbreite der Anwendungen von Fallstudien in FTI Evaluationen ist groß. Sie reicht von der einfachen Illustration von Best Practice Cases, über die triangulative Kombination von Fallstudien mit anderen qualitativen und quantitativen Forschungsmethoden, bis hin zu komplexeren Kontextanalysen und der Identifikation und Abschätzung einzelner Wirkungsketten.

Im vorliegenden Beitrag wird den Fragen nachgegangen, welche Zwecke mit Fallstudien verfolgt werden, welche Arten von Fallstudien in welchen Evaluationskontexten eingesetzt werden, welche Aspekte bei der Falldefinition und –auswahl beachtenswert sind, und unter welchen Bedingungen Fallstudien welchen Nutzen in Evaluationen bringen können.

Methodisch basiert der Beitrag auf einer Analyse einschlägiger Literatur und wird ergänzt um ein Screening von FTI Evaluationen, die Fallstudien inkludieren.¹

Die Conclusio ist, dass die Anwendung von Fallstudien fallweise unter ihrem Potenzial bleibt. Wir argumentieren dafür, Fallstudien im Rahmen von Evaluationen systematischer einzusetzen, den Zweck von Fallstudien im Kontext der Evaluation genauer zu definieren, und anhand eines FTI Fallstudien-Repositoriums deren interne und externe Validität sowie Reliabilität zu erhöhen.

KEYWORDS:

Fallstudien, Evaluation, FTI Politik, Multi-Methoden-Einsatz

Fallstudien sind in der Evaluation von FTI Forschungsförderung ein relativ häufig genutztes Instrument der empirischen Sozialforschung. Das INNO Appraisal Projekt identifizierte zum Beispiel bei 41 % von 216 europäischen FTI Evaluationen die Anwendung von Fallbeispielen und auch innerhalb der von der EU-GD Forschung und Innovation in Auftrag

gegebenen Evaluationen der letzten Jahre wurden in 49 % (von 21 Studien in 2011) bzw. 60 % (von 32 Studien in 2013) Fallstudien angewendet (Edler et al. 2010: iv, DG RTD 2011: 11; DG RTD 2013: 70).

TYPEN UND ZWECKE VON FALLSTUDIEN

Eine Fallstudie ist ein Instrument zur empirischen Untersuchung eines Phänomens in seinem realweltlichen Kontext, insbesondere wenn die Grenzen zwischen Phänomen und Kontext nicht deutlich zu erkennen sind. Im Unterschied zu Fallstudien in einem „reinen“ Forschungskontext gilt es bei Fallstudien im Rahmen von Evaluationen nicht unbedingt Hypothesen als Ausgangspunkt zu wählen, sondern Fallstudien als Strategie zu verwenden, um das Evaluandum (z.B. eine geförderte Organisation) im Detail zu studieren - in der Regel ohne besonders gewichtige Vorannahmen. Gerade in Evaluationen steht bei der Anwendung von Fallstudien, das Lernen über komplexe Sachverhalte im Vordergrund. Dies basiert auf tiefgehenden Analysen und einem dadurch erzielten umfassenden Verständnis des Sachverhalts und dessen Kontext (Yin 2003).

In vielen Evaluationen werden Fallstudien als Methode (neben anderen) dargestellt; genaugenommen handelt es sich bei Fallstudien aber um eine umfassende Forschungsstrategie, die sich einzelner Methoden als Techniken der Informations- und Datenerhebung bedient, um Antworten auf Forschungs- bzw. Evaluationsfragen zu finden. Demnach soll eine Fallstudie ein maßgeschneidertes Design aufweisen, das auf die jeweiligen Evaluationsziele bzw. -fragenstellungen und empirischen Rahmenbedingungen abgestimmt ist, und mehrere Informationsquellen und Methoden inkludiert. Insbesondere der Anspruch, den relevanten Kontext mit zu erfassen, verbietet es nur eine Methode einzusetzen. Daher ist eine Fallstudie per se multi-methodisch anzulegen und im Prinzip offen für alle Methoden der empirischen Sozialforschung (Lamnek 1995: 5). Die Entscheidung, Fallstudien als Forschungsstrategie einzusetzen,

¹ Zunächst wurden dazu alle auf der fteval-Plattform veröffentlichten Evaluierungen seit 2009 mit Blick auf die Anwendung von Fallstudien hin gescreent und anhand eines Kriterienrasters systematisiert. Über diese 16 (von 41) Evaluierungen hinaus wurden auch noch zwei interessant erscheinende Beispiele aus Deutschland und eine Evaluierung auf EU-Ebene berücksichtigt, die einen besonders umfassenden Fallstudienansatz inkludierte. Die empirische Basis an Evaluierungen (n=19) ist der Literaturliste zu entnehmen.

und die genauere Bestimmung des Zweckes der Fallstudien ergibt sich meist aus der Art der Evaluation und den damit verbundenen Evaluationszielen und –schwerpunkten (Stern 2002: 173).

In der einschlägigen Literatur werden drei Hauptkategorien („descriptive“, „exploratory“ und „combined/cumulative“) und sechs Fallstudientypen unterschieden, die sich in der Forschungs- und Evaluationsrealität durchaus überlappen (Yin 2003; Morra/Friedlander 1990; GAO 1990). Im Wesentlichen wird zwischen deskriptiven und erklärenden Typen unterschieden.

Deskriptive Fallstudien: Die deskriptiven Fallstudien („descriptive case studies“) gliedern sich wiederum in „illustrative“, „explorative“ und „critical instance“ Fallstudien. Die *illustrative Fallstudie* beschreibt in erster Linie einen Fall oder eine begrenzte Anzahl von Fällen und hat den Zweck eines Beispiels für andere, bereits erhobene Information über ein Programm, Projekt, etc. Dabei dient die Fallstudie der besseren Veranschaulichung und Untermauerung von Evidenz aus anderen (methodischen) Quellen. Die *explorative Fallstudie* soll hauptsächlich als hypothesengenerierender Einstieg in ein Feld dienen, in dem die Informationen vorab nur in geringem Ausmaß vorhanden sind. Die Darstellung einer explorativen Fallstudie ist meist zweitrangig, da ihr Zweck hauptsächlich darin besteht, eine fundierte Basis für nachfolgende Evaluationsarbeiten zu bilden. Die *„critical instance“ Fallstudie* untersucht ein ganz bestimmtes Spezifikum, das im Evaluationskontext von entscheidender Bedeutung ist und hat den Zweck, eine bestimmte Behauptung bzw. Annahme über einen Programm- oder Projektaspekt kritisch zu testen.

Erklärende Fallstudien: Erklärende Fallstudien („explanatory case studies“) sind weniger fokussiert auf einen bestimmten Sachverhalt, sondern haben den Zweck, Zusammenhänge einzelner (Programm)elemente zu erklären. Im Wesentlichen wird zwischen zwei Typen unterschieden: „programme implementation“ Fallstudien und „programme effects“ Fallstudien. Die *„programme implementation“ Fallstudie* hat den Zweck, die Programmdurchführung im Sinne der operativen Prozesse – oft an unterschiedlichen Orten und überwiegend normativ - im Detail zu untersuchen. Evaluationsfragen zur Umsetzung, insbesondere zu Effektivität und Effizienz, stehen dabei im Mittelpunkt. Die *„programme effects“ Fallstudie* analysiert die Auswirkungen von Programmen, Projekten, etc., - positiv und negativ, intendiert und nicht-intendiert - im Sinne kausaler Zusammenhänge zwischen Ursachen und Wirkungen. Es gilt dabei besonders den Kontext zu verstehen und daher zu prüfen, ob und inwiefern die Wirkungen aus der Umsetzung eines Programms resultieren, wodurch hier durchaus Überschneidungen mit der „programme implemen-

tion“ Fallstudie gegeben sind. Ein wesentliches Merkmal dieses Fallstudientyps ist ein oder auch mehrere analytische Schritte, die innerhalb der Fallstudie auf Basis der erhobenen Informationen getätigt werden.

Kumulative Fallstudien: In „cumulative case studies“ werden Ergebnisse aus einer Mehrzahl von Fallstudien (deskriptiv, normativ oder Ursache-Wirkung) zusammengebracht, um Evaluationsfragen auf einem höheren Aggregationsniveau zu beantworten. Oft wird bei kumulativen Fallstudien auch die historische Perspektive miteinbezogen.

WELCHE ART VON FALLSTUDIEN IN WELCHEN EVALUATIONSARTEN?

Ein Blick auf durchgeführte FTI Evaluationen der letzten Jahre in Österreich und auf einzelne interessante Beispiele darüber hinaus zeigt, dass fast 40 % der Evaluationen, die seit 2009 über die Österreichische Plattform für Forschungs- und Technologiepolitik (fteval) öffentlich zugänglich sind, Fallstudien inkludierten. Fallstudien werden in allen Arten der Evaluation (Ex-ante, Zwischen- und Ex-post sowie begleitende Evaluation) verwendet. Von den untersuchten FTI Evaluationen mit Fallstudien sind 53 % Interims- und 37 % Ex-post oder Endevaluationen. Auch im Rahmen des INNO-Appraisal-Projekts (Edler et al. 2010) wurden Fallstudien vor allem in Interims- und Ex-post Evaluationen identifiziert.

In sieben der gescreenten Berichte haben die Fallstudien einen klar deskriptiven bzw. illustrativen Charakter; elf Evaluationen weisen Fallstudien auf, die entsprechende Zusammenhänge und Kontexte in erklärender Art und Weise aufzeigen. Diese Fallstudien weisen erkennbare Analysearbeiten auf, wenn auch unterschiedlichster Ausprägung, und in fünf dieser Evaluationen wird im Rahmen der erklärenden Fallstudien auch Bezug auf (Programm)Implementierungsaspekte genommen. In einer Evaluation wurde mit den Fallstudien die Illustrierung einiger weniger spezifischer Problemstellungen bezweckt (critical instance). Explorierende Fallstudien beinhaltete keine der untersuchten Evaluationen. Dies ist nicht weiter verwunderlich, da diese Art von Fallstudien i.d.R. nicht in Berichten dokumentiert wird. Es zeigt sich, dass deskriptive Fallstudien eher in Interims- und begleitenden Evaluationen eingesetzt wurden, während Fallstudien des Typs „programme effects“ sowohl in Interims- wie auch in Ex-post bzw. Endevaluationen zu finden sind.

	Ex-ante	Interim	Ex-post/End	Begleitend
deskriptive Fallstudien		5	1	2
erklärende Fallstudien		5	6	
kumulative Fallstudien				

Tabelle 1: Typen von Fallstudien in untersuchten Evaluationen (2009-2014)

Quelle: <http://www.fteval.at/> (4.11.2014), Evaluationsberichte mit Fallstudien (2009-2014). Eigene Auswertung.

Grundsätzlich ist festzuhalten, dass die Art des Fallstudientyps weniger allein durch die Art der Evaluation bedingt wird, sondern vielmehr von den Evaluationszielen und –schwerpunkten abhängt. Das Aufzeigen von komplexen kausalen Zusammenhängen („programme effects“ Fallstudien) kann nicht nur am Ende einer zu evaluierenden Maßnahme Sinn machen, sondern beispielsweise auch im Rahmen von Programmbegleitungen, um spezifische Mechanismen in „real time“ beobachten zu können und Ergebnisse entsprechend in die Politikgestaltung einfließen zu lassen. Andererseits können - gerade im FTI-Bereich relevant - zeitlich länger dauernde Wirkungsketten besser aus einer ex-post Perspektive beurteilt werden. Dennoch kann es nützlich sein, begleitend oder in Interimsevaluationen mit Fallstudien eine Basis für Impact-Fallstudien zu legen, die im Rahmen einer End- oder Ex-post Evaluation weitergeführt werden – eine Vorgehensweise, die in den untersuchten Fällen nicht gefunden wurde.

WAS IST DER FALL IN EINER FALLSTUDIE? DESIGN UND AUSWAHLFRAGEN

Die Art der Evaluationsfragen sowie der zu evaluierenden Maßnahme ist entscheidend für die Definition der Analyseeinheit (der konkrete Fall). Robert Yin unterscheidet zwischen einem single-case Design und einem multiple-case Design (Yin 2003: 39ff, und Abbildung 1). Im Rahmen des single-case Designs stehen ein Fall und sein Kontext im Zentrum der Analyse (holistic-single), wobei den Fall auch mehrere Einheiten bilden können (embedded-single). Dem gegenüber werden bei dem multiple-case Design mehrere Fälle analysiert. Auch dabei können ein Fall und sein Kontext im Zentrum stehen (holistic-multiple), oder aber es können mehrere Einheiten den Fall in seinem Kontext bilden (embedded-multiple).

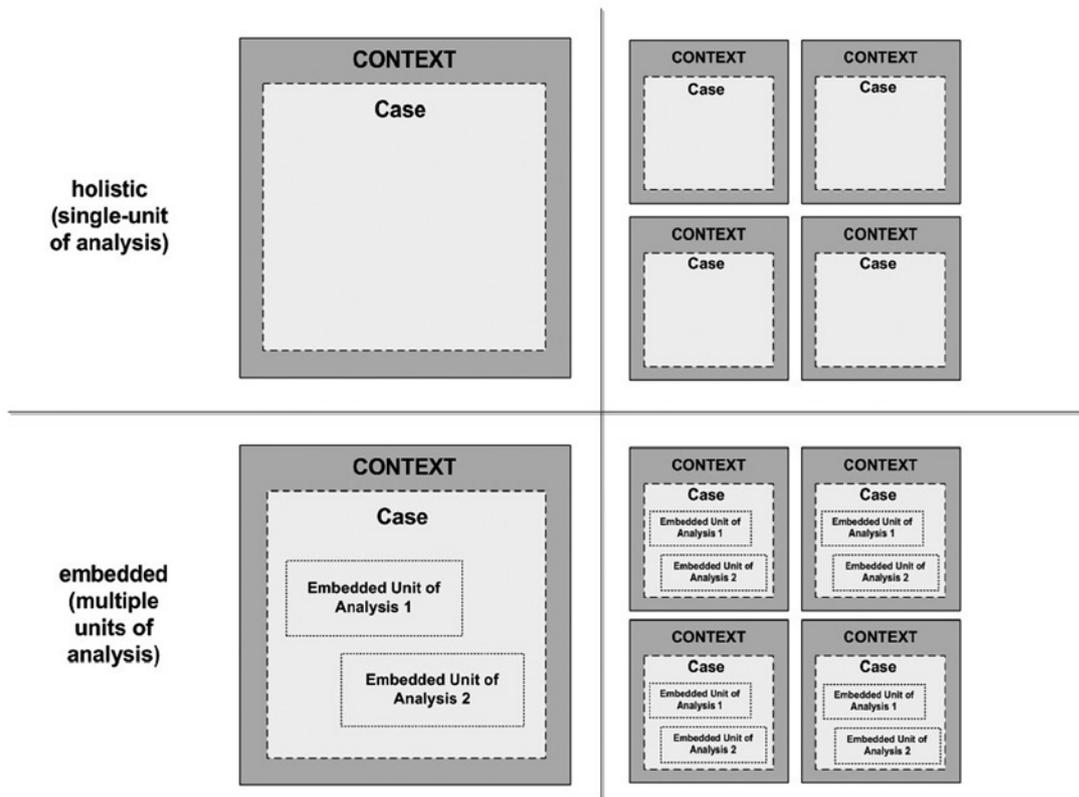


Abbildung 1: Grundtypen für das Design von Fallstudien
 Quelle: Yin (2003) nach COSMOS Corporation

Im FTI Evaluationskontext sind Fälle zumeist Projekte in Verbindung mit (ausgewählten) geförderten Organisationen, aber auch geförderte Individuen, Länder/Regionen oder spezifische Maßnahmen. Die Fallstudien-Designs sind dabei immer multiple-case Designs und vielfach bilden auch mehrere Analyseeinheiten einen Fall: Beispielsweise steht ein Projekt unter Berücksichtigung der teilnehmenden Organisationen im Zentrum der Fallstudie (siehe Abbildung 2, „Fall 1“) oder eine (ausgewählte) Organisation mit zwei oder mehr Partnerorganisationen in einem gemeinsamen Projekt bilden einen Fall (siehe Abbildung 2, „Fall 2“). Sehr oft wird nicht nur ein Projekt sondern vielmehr eine Organisation im Projektkontext als Fall analysiert, der wiederum seinen eigenen Kontext hat. Beispielsweise kann eine einzelne Organisation als eine Analyseeinheit, und ein Projekt, an dem diese Organisation teilgenommen hat, als eine

zweite Analyseeinheit definiert sein, die gemeinsam einen Fall darstellen. Damit steht die Projektteilnahme der Organisation im Zentrum der Fallstudie. Beide Analyseeinheiten haben wiederum für die Fallstudie äußerst relevanten Kontext. Die Organisation mit seinen generellen FTI-Aktivitäten, seiner Branchenzugehörigkeit etc. und das Projekt in seinem Themenbereich, mit weiteren Projektteilnehmern, etc.

Auch kann eine Fallstudie ein ganzes Projekt mit allen teilnehmenden Organisationen beinhalten (siehe Abbildung 2, „Fall 3“). Darüber hinaus gibt es in einem derartigen Mikrokosmos noch zahlreiche andere Spielarten, den passenden Fall zu definieren. Wesentlich ist als Ausgangspunkt der Analysen eine bewußte Klärung, welche Einheiten den Fall bilden und darüber hinaus, was relevanter Kontext ist.

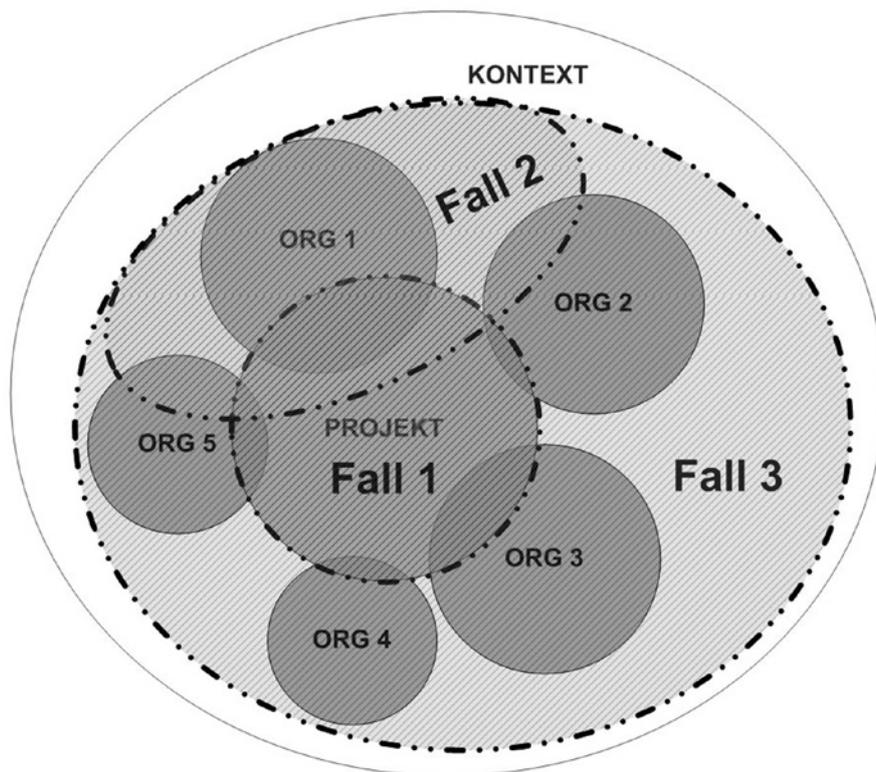


Abbildung 2: Was ist mein Fall? Projekt vs. Organisation
Quelle: KMU Forschung Austria

Unter den untersuchten FTI Evaluationen finden sich im Wesentlichen drei Typen an Fällen, die jeweils die Basis der Fallstudie bilden. In fünf Evaluationen bilden Länder bzw. Maßnahmen in anderen Ländern (i.S. eines internationalen Vergleichs) die Fälle, in drei Evaluationen sind es Individuen (Personenförderungen) und in elf Evaluationen wurden Projekte in Verbindung mit Organisationen als Fälle untersucht. Bei letzterem standen abwechselnd Projekte und beteiligte Organisationen im Vordergrund. Die Fallstudien über Länder bzw. Maßnahmen in anderen Ländern im Rahmen internationaler Vergleiche waren naturgemäß alle von deskriptivem Charakter, jene mit Individuen als zentrale Untersuchungseinheiten hatten alle erklärenden Charakter und die Fallstudien über Projekte in Verbindung mit Organisationen sind überwiegend dem erklärenden Fallstudientypus zuzurechnen.

Die Auswahl der Fälle ist eng verbunden mit den Evaluationszielen und entsprechenden Fragestellungen. Diese Auswahl erfolgt in der Regel zweckbestimmt, das heißt sie richtet sich nach dem zentralen Erkenntnisinteresse. Ist es etwa im Rahmen einer Evaluation besonders wichtig, das Verhalten einer oder unterschiedlicher Organisationen zu verstehen (Zielgruppe/n)? Oder ist es sinnvoll anhand von außerhalb der Norm gelagerten Fällen Unterschiede herauszuarbeiten (Kontrastierung)? Des Weiteren kann eine zentrale Fragestellung sein, warum einige Projekte besonders erfolgreich sind und warum andere nicht oder weniger gut funktionieren (Performance)? Umgekehrt kann aber auch von Interesse sein, was typischerweise warum in Projekten passiert (typische Fälle), etc.

WAS KÖNNEN FALLSTUDIEN LEISTEN UND WORAUF IST ZU ACHTEN?

Fallstudien sind eine gelegentlich unterschätzte Strategie für die Bewältigung und vertiefende Analyse komplexer Sachverhalte und Situationen, insbesondere wenn quantitative Daten knapp oder nicht verfügbar sind und Antworten auf "Warum" und "Wie" Fragen ebenso wichtig oder noch wichtiger sind, als Antworten auf „Wieviel“ Fragen alleine. Auch kann bei Fallstudien relativ flexibel und offen gegenüber unvorhergesehenen Zusammenhängen reagiert werden und es ist von zentraler Bedeutung, mit einer/mehreren Untersuchungseinheit(en) zu interagieren sowie den Kontext, in dem sie sich bewegen und wichtige Prozesse stattfinden, miteinzubeziehen.

FALLSTUDIEN ZUR ERKLÄRUNG VON WIRKUNGSZUSAMMENHÄNGEN

Die meisten Ausschreibungskonzepte zu FTI Evaluationen - naheliegender wenn es sich um Endevaluationen handelt, aber auch bei Interimsevaluationen - stellen Fragen nach Wirkungen und Wirkungsweisen der zu evaluierenden Maßnahmen. Die Erklärung von Wirkungszusammenhängen mittels Kausalketten, oft über einen gewissen Zeitraum hinweg – begleitend oder retrospektiv -, ist in Evaluationen besonders von Interesse, und mit Fallstudien gut erreichbar.

Um die Wirkungen einer Intervention festzustellen, ist es notwendig die Implementierung der Intervention - unter Einbezug der Kontexte - Schritt für Schritt nachzuverfolgen, festzuhalten („tracing and mapping“) und darüber hinaus die wesentlichen Einflussfaktoren herauszuarbeiten. Dazu empfiehlt es sich in zwei Phasen vorzugehen (Bjurulf 2012: 60): zunächst gilt es die Interventionslogik im Sinne der intendierten Wirkungen der Intervention zu rekonstruieren und damit zu zeigen, wie die Intervention schrittweise intendiert wirken soll. Diese theoretische Wirkungskette kann in einer zweiten Phase als Hypothese empirisch getestet werden. Die Nachzeichnung einer Wirkungskette dient zur Klärung der Richtung und Stärke von kausalen Abläufen über einen gewissen Zeitraum hinweg und kann „windows of opportunity“ und zentrale Einflussfaktoren für Wirkungen identifizieren.

Um fundierte Aussagen über Wirkungen treffen zu können und vor allem um Wirkungszusammenhänge zu verstehen, sind Fallstudien als Strategie, die idealerweise mehrere Methoden vereint und damit auch den Betroffenen (beispielsweise eine geförderte Organisationen) involviert, äußerst nützlich. Analytisch quantitative Methoden können, je nach Definition, entweder in solche Fallstudien integriert werden, oder aber auch können Fallstudien und analytisch quantitative Methoden nebeneinander stehen und sich auf Evaluationsebene ergänzen.

MULTI-METHODEN-EINSATZ

Die Analyse qualitativer Daten gewährt Tiefeneinblicke, beispielsweise in direkte und indirekte Effekte einer Maßnahme und vor allem in diesbezügliche kausale Zusammenhänge, die Quantifizierungen alleine niemals freilegen können. Doch gerade diese „Weichheit“ der qualitativen Daten ist es, welche die Fallstudie als Forschungsstrategie auch

angreifbar macht. Zu den gängigen Vorbehalten, die sich oft gegen Fallstudien als Evaluationsstrategie richten, zählen eine geringere Objektivität, Quantifizierbarkeit, repräsentative Aussagekraft, Generalisierbarkeit oder Robustheit als andere Ansätze, wie beispielsweise rein quantitative Ansätze über große Stichproben in Form von standardisierten Befragungen oder Analysen von Sekundärdaten.

Der Prozess der Informationserhebung im Rahmen von Fallstudien gestaltet sich deutlich komplexer als in anderen Forschungsansätzen. Vielfach wird in der Fallstudienforschung für einen konsistenten Multi-Methodeneinsatz plädiert (Yin 2003, Terluin/Berkhout 2011, Stern 2002, Morra/Friedlander 1990), gerade wenn eine Evaluation schwerpunktmäßig oder einzig auf einem Fallstudienansatz basiert. Dazu sollen unterschiedliche Forschungsmethoden (Interviews unterschiedlicher Art, standardisierte Befragungen, Dokumenten- und Inhaltsanalyse, Sekundärstatistiken – durchaus auch mittels ökonomischer Methoden analysiert, teilnehmende Beobachtung, etc.) angewendet und die Ergebnisse in Form einer interpretativen Synthese trianguliert werden. Die phasenweise Kombination der Evidenzsammlung aus mehreren Quellen – beispielsweise zunächst Sekundärdaten (was und wieviel?), dann Tiefeninterviews (wie und warum?) - soll die Feststellung der tatsächlich erreichten Wirkungen am besten ermöglichen. Dieser Prozess ist in vielen Fällen sinnvollerweise iterativ gestaltet.

Wenngleich ein Methodenmix und die Triangulation in FTI Evaluationen zu einem Standard geworden sind, haben Fallstudien oftmals ein eher geringes Gewicht im gesamten Ansatz der Evaluation und bleiben dementsprechend illustrative „Zusatzgeschichten“. In neun der untersuchten FTI Evaluationen haben Fallstudien ein eher geringes Gewicht (die Mehrheit dieser Evaluationen weist Fallstudien des deskriptiven Typus auf); in weiteren neun Evaluationen haben die durchgeführten Fallstudien ein merkliches Gewicht in der Evaluation (diese Evaluationen beinhalten ausschließlich Fallstudien des erklärenden Typus) und nur eine Evaluation baut gänzlich auf einem Fallstudienansatz auf.

Der systematische Einsatz von mehreren, aufeinander abgestimmten Methoden zur Erstellung von Fallstudien ist dann unablässig, wenn die Fallstudien über einen deskriptiv-illustrativen Charakter hinausgehen, also stärker erklären und damit in der Regel Wirkungswege und –ketten rekonstruieren und offenlegen sollen. Unter den untersuchten Evaluationen, die erklärende Fallstudien durchführten, wurden jeweils mindestens zwei und mehrheitlich drei Methoden – darunter immer Dokumentenanalyse und Interviews bzw. als dritte Methode meist die Analyse von Monitoring- oder Sekundärdaten – als Informationsquellen herangezogen. Unter diesen war die Triangulation innerhalb der Fallstudien immer ausreichend nachvollziehbar und – mit einer Ausnahme - auch auf Gesamtevaluationsebene war dem so. In zwei Evaluationen wurde neben Dokumentenanalyse und Interviews auf bibliometrische Analysen zurückgegriffen.

Je stärker ein Fallstudienansatz einen Schwerpunkt in der Evaluation bildet, desto differenzierter muss auch der Methodeneinsatz erfolgen, was sich nicht zuletzt auch durch den Umstand ergibt, dass die einzelnen Methoden in diesen Fällen auch exklusiv für die Fallstudien eingesetzt werden. Die Wahl eines Fallstudienansatzes – unabhängig in welcher Ausprägung – ist auch immer von der Art und dem Zeitpunkt der Evaluation und den sich dadurch ergebenden Evaluationsfragen abhängig.

Nahezu alle untersuchten Ex-post- bzw. EndEvaluationen wie auch die Hälfte aller untersuchten Interimsevaluationen haben Fallstudien des erklärenden Typus als deutlichen Schwerpunkt im Gesamtevaluationskonzept verwendet.

QUALITÄT SICHERSTELLEN

Die Qualität von Fallstudien kann nach *vier Kriterien* getestet werden, die in der empirischen Sozialforschung breite Anwendung erfahren und für Fallstudien im Speziellen mit der Beachtung von gewissen Aspekten bzw. Taktiken in Design, Erhebung und Analyse verstärkt werden (Yin 2003: 33ff). Für den Test der *Konstrukt-Validität* sind operationalisierte Kriterien für die Dateninterpretation zu entwickeln, um eine subjektive Färbung der Bewertung von erhobenen Informationen bestmöglich zu umgehen. Die Konstrukt-Validität kann durch den Einsatz mehrerer Informations- und Datenquellen und durch die Führung einer Beweiskette erhöht werden. Die *interne Validität* ist hauptsächlich bei erklärenden Fallstudien gefragt, in denen kausale Zusammenhänge untersucht werden. Schlußfolgerungen über angenommene Zusammenhänge müssen vollständig transparent sein; rivalisierende Erklärungen müssen abgewogen und ausgeschlossen werden können. Interne Validität in der Analyse der erhobenen Daten kann am besten mittels der Verwendung von Logic Models, der Untersuchung von rivalisierenden Erklärungen und einem „pattern matching“² erreicht werden.

Der dritte Qualitätstest zur Feststellung der *externen Validität*, beschäftigt sich mit der Frage, ob die Ergebnisse auf Evaluationsebene über die Studiengrenzen hinaus generalisierbar sind, was in der Fallstudienforschung als problematisch angesehen wird. Schlußfolgerungen aus einer Fallstudie können erhärtet werden, indem ein multiple-case Design angewendet wird und die Replikation der Ergebnisse im Rahmen anderer Fallstudien gelingt. Schließlich kann eine Fallstudie auch auf *Reliabilität* geprüft werden; dieses Kriterium ist erfüllt, wenn man im Rahmen einer anderen Evaluation bei erneuter Durchführung derselben Fallstudie unter Anwendung derselben Vorgehensweise zu denselben Ergebnissen kommt. Geeignete Strategien um Fallstudien annähernd reliabel zu machen sind eine gute Dokumentation der Vorgehensweise und Erhebungen (Protokolle und Datensatz).

Unter diesen Testmöglichkeiten sind gerade die Feststellung der externen Validität und der Reliabilität in der Evaluationsrealität schwierig und nicht von zentraler Bedeutung. Fallstudien müssen demnach nicht rigoros all diese Qualitätstests „bestehen“. Finden sich beispielsweise keine vergleichbaren Evaluationen mit vergleichbaren Fallstudien, so kann sich der/die EvaluatorIn nur mittels gewisser Taktiken, wie oben ausgeführt, einer externen Validität nähern. Auch wird in der Evaluationsrealität niemand dieselbe Fallstudie im selben Kontext bzw. unter Anwendung derselben Vorgehensweise erneut erstellen; eine genaue Dokumentation ist jedoch unumgänglich.

GRENZEN IN DER ANWENDUNG VON FALLSTUDIEN IN DER EVALUATIONSREALITÄT

Einige der in der Fallstudienforschung vorgeschlagenen Regeln für deren Anwendung sind in der praktischen Umsetzung bei Evaluationen oft nicht wissenschaftlich exakt umsetzbar – sei es aufgrund zeitlicher oder budgetärer Restriktionen, oder aufgrund spezifischer Umstände, wie z.B. ein begrenzter Zugang zu Analyseeinheiten oder nicht verfügbare Daten. Dennoch kann sich der/die EvaluatorIn, wie oben am Beispiel der Qualitätsprüfungen bereits kurz ausgeführt, einer wissenschaftlich korrekten Herangehensweise möglichst annähern und damit die interne Validität nicht zu stark kompromittieren. Die konkrete Ausgestaltung, und damit auch die Tiefe der Analyse hängt nicht nur von der Natur des Untersuchungsobjekts ab, sondern auch vom Ziel, der Komplexität und der Verortung der Evaluation im politischen Prozess. Damit hängen auch die Definition, die Auswahl und die Anzahl der Fälle sowie die einzusetzenden Erhebungsmethoden zusammen.

Insbesondere Fallstudien, die auf die Analyse von Wirkungszusammenhängen abstellen, benötigen ein stringentes Design und die Sensitivität gegenüber alternativen Erklärungsmustern. Die Umsetzung von Forschungsstrategien mit kausalem Anspruch in einem komplexen Umfeld erfordert vor allem konzeptionelles Wissen in der Sache, und Erfahrung mit der Umsetzung dieser Forschungsstrategie. Sind Evaluationen nicht mit ausreichenden finanziellen und zeitlichen Ressourcen ausgestattet und werden sie nicht von erfahrenen ForscherInnen durchgeführt, leidet i.d.R. die Qualität der Fallstudien, und es stellt sich die Frage nach deren Mehrwert.

Um die etwas begrenzenden Rahmenbedingungen in der Evaluationsrealität abschwächen zu können, wäre es überlegenswert, ein Fallstudien-Repository anzulegen in dem Fallstudien im FTI-Bereich (nicht öffentlich, im Rahmen der fteval) dokumentiert werden. Hierzu hätten EvaluatorInnen Zugriff – einerseits um die eigene Fallstudienvorbereitung zu unterstützen, bzw. auch um die eigenen Ergebnisse mit anderen Fallstudienresultaten auf ihre externe Validität und Reliabilität prüfen zu können.

ZUSAMMENFASSUNG UND FAZIT

In zahlreichen Evaluationen werden Fallstudien als eine neben anderen Methoden dargestellt; genaugenommen handelt es sich bei Fallstudien jedoch um eine umfassende Forschungsstrategie, die ein maßgeschneidertes Design aufweist, das auf die jeweiligen Evaluationsziele bzw. -fragestellungen und empirischen Rahmenbedingungen abgestimmt ist, und mehrere Informationsquellen und Methoden inkludiert.

2

Beim „pattern matching“ werden die aus empirischen Erhebungen resultierenden Muster bzw. Wirkungszusammenhänge mit vorangegangenen Mustern verglichen um Gemeinsamkeiten bzw. Unterschiede hervorzuheben und daraus neue Erkenntnisse abzuleiten. Wenn die Muster übereinstimmen, können die Ergebnisse die interne Validität stärken. Weitere Spielarten sind denkbar, wie z.B. zur Etablierung eines Kontrollgruppenansatzes anhand des Vergleichs von Fällen.

Fallstudien sollten im Evaluationskontext der vertiefenden Analyse komplexer Sachverhalte dienen, insbesondere wenn Antworten auf "Warum" und "Wie" Fragen ebenso wichtig oder noch wichtiger sind, als Antworten auf „Wieviel“ Fragen. Dabei sind der Einbezug von Kontext und die Interaktion mit den Untersuchungseinheiten von zentraler Bedeutung.

Im Wesentlichen wird zwischen deskriptiven und erklärenden Fallstudien unterschieden. Während deskriptive Fallstudien der Illustration eines Phänomens dienen, zielen erklärende Fallstudien darauf ab, Wirkungszusammenhänge mittels Kausalketten nachzuverfolgen und analytisch herauszuarbeiten.

Gerade in der FTI-Politik, mit typischerweise zeitlich länger dauernden Wirkungsketten, werden Wirkungen im Spektrum der Outputadditionalität besser aus einer ex-post Perspektive analysiert. Hier können Fallstudien Fragen nach der Hebelwirkung von Förderung in der Genese einer ForscherInnenkarriere, einer Technologie, eines Produkts, etc. beantworten.

Im Rahmen von Zwischenevaluationen ist zu beachten, dass nur ein bestimmter Typus von kurz- bis mittelfristigen Effekten - hauptsächlich im Spektrum der Verhaltensadditionalität - während der Programmumsetzung abgebildet werden kann. Um den Kontext und spezifische Mechanismen in „real time“ beobachten zu können, bieten sich Fallstudien auch im Rahmen von begleitenden Evaluationen an. Naheliegende Anwendungsgebiete sind z.B. innovative Formate von Programmen, in denen der Kontext sowie die Entwicklung von Projekten ausgeleuchtet werden soll und Ergebnisse unmittelbar in die Politikgestaltung einfließen sollen.

Darüber hinaus kann es nützlich sein, begleitend oder in Interimsevaluationen mit Fallstudien eine Basis für Impact-Fallstudien zu legen, die im Rahmen einer Endevaluation durchgeführt werden; eine Vorgehensweise, die in den untersuchten Evaluationen bedauerlicherweise nicht gefunden wurde, da diese in Impact-Studien durchaus weitergeführt werden könnten.

Im Zuge der Recherche auf der fteval-Webpage wurden in begleitenden und Interimsevaluationen sowohl deskriptive als auch erklärende Fallstudientypen gefunden, wobei letztere die Mehrheit stellen und, wenn auch in unterschiedlicher Ausprägung, erkennbare Analytik aufweisen. Nur eine Evaluation inkludierte „critical instance“ Fallstudien; ein interessanter Subtypus der deskriptiven Fallstudien, die einen bestimmten Teilaspekt eines Phänomens untersuchen. Im Evaluationskontext könnte man damit spezifische Annahmen über z.B. die Hebelwirkung von (nicht) erfolgreichen Projektumsetzungen durchleuchten. Dieser Fallstudientypus könnte zukünftig in so mancher FTI Evaluation einen Mehrwert darstellen.

In Bezug auf das Fallstudiendesign ist zu klären, welche Einheiten sinnvoller Weise einen Fall bilden und was relevanter Kontext ist. In FTI Evaluationen bilden zumeist Projekte in Verbindung mit ausgewählten, geförderten Organisationen die Fälle; aber es finden sich mitunter auch geförderte Individuen, Länder/Regionen oder spezifische Maßnahmen. Die Fallstudiendesigns sind dabei immer multiple-case Designs; d.h. es werden mehrere Fälle analysiert, und vielfach bilden auch mehrere Analyseeinheiten einen Fall.

Der Prozess der Informationserhebung im Rahmen von Fallstudien gestaltet sich deutlich komplexer als in manch anderen Forschungsansätzen. Vielfach wird in der Fallstudienforschung für einen konsistenten Multi-Methodeneinsatz plädiert. Je stärker ein Fallstudienansatz einen Schwerpunkt in der Evaluation bildet, desto differenzierter muss auch der Methodeneinsatz erfolgen. Die unterschiedlichen Methoden werden im gesamten Prozess phasenweise oder auch iterativ eingesetzt. Wenngleich ein Methodenmix und die Triangulation in FTI Evaluationen zu einem Standard geworden sind, haben Fallstudien oft ein geringes Gewicht im gesamten Ansatz der Evaluationen und bleiben dementsprechend illustrative Zusatzgeschichten. Um Fallstudien in Evaluationen mehr Gewicht geben zu können und um qualitativ hochwertige Ergebnisse zu erzielen, die einen dezidierten Mehrwert in der Evaluation darstellen, bedarf es einer passenden Ausstattung an finanziellen, zeitlichen und Humanressourcen.

LITERATUR

Alt, R., Berrer, H., Borrmann, J., et al., (2012): Nutzen-, Programm- und Systemevaluierung der Christian Doppler Forschungsgesellschaft. Wien.

Bjurulf Staffan, Vedung Evert, Larsson, C.G. (2012): A triangulation approach to impact evaluation. *Evaluation* 19(1) 56-73.

Daimer, S., Berghäuser, H., Braun, A., et al., (2014): Begleitende Evaluierung der Fördermaßnahme „Validierung des Innovationspotenzials wissenschaftlicher Forschung – VIP“. Executive Summary. Berlin.

Dinges, M., Leitner, K., Zahradnik, G., (2014): Evaluierung des Programmes FHplus. Endbericht. Wien

Ecker, B., Kottmann, A., Meyer, S., (2014): Evaluation of the FWF Doctoral Programme (DK Programme). Research Report. Wien.

Edler, J., (2010): INNO-Appraisal - Perspectives on Evaluation and Monitoring. Final Report. Brussels, iv.

Elk van K., et al. (2014), Performance of SMEs within FP7. An interim evaluation of FP7 components. Brussels.

European Commission (2013): DG RTD Annual Report on Programme Evaluation Activities 2011.

European Commission (2014): DG RTD Annual Report on Programme Evaluation Activities 2013.

Fischl, I., Gustavsson, M., Romanow, B., Ruhland, S. (2010): "Strategic impact, no revolution." Ex-post evaluation of the thematic priority Nanotechnology and nanosciences, knowledge-based multifunctional materials and new production processes and devices. Strategic Level. Brussels.

GAO-United States General Accounting Office. Program Evaluation Methodology Division (1990): Case Study Evaluations. Washington D.C. Geyer, A., Tiefenthaler, B., (2011): Programmevaluierung „Headquarter-Strategy“. Endbericht. Wien.

Heckl, E., Sheikh, S., Wolf, L. (2014): Zwischenevaluierung des Förderschwerpunkts Talente. Endbericht. Wien.

Kind, S., Kaufmann, P., Ehrenberg-Silies, S., Fischl, I., Hannicke, S., Hoppe, U., Wolf, L. (2014): Endbericht zur Erfolgskontrolle des Programms zur Förderung der Industriellen Gemeinschaftsforschung (IGF). Berlin.

Knoll, M., Illge, L., Handke, V., et al. (2014): Evaluation der Projektförderung des Bundesministeriums für Wirtschaft und Technologie in der Energieforschung, Fachbereich Energieeffizienz in Industrie, Gewerbe, Handel und Dienstleistung (IGHD) im Rahmen des 5. Energieforschungsprogramms. Endbericht. Berlin.

Lähteenmäki-Smith, K., Halme, K., Lemola, T., et al. (2013): "Licence to SHOK?" – External Evaluation of the Strategic Centres for Science, Technology and Innovation.

Lamnek, S. (1995): Qualitative Sozialforschung. Bd. 2 Methoden und Techniken. 3. Auflage. Psychologie Verlags Union, Weinheim.

Laudel, G. (2013): An In-depth Case Study of Selected WWTF Impacts. Impact Evaluation 2013/14. Berlin.

Morra, L. G., Friedlander, A. C. (1999): Case Study Evaluations. The World Bank. Washington D.C.

Pfirmsmann, O., Heinrich, S., Riesenberger, D., et al. (2012): Evaluierung des österreichischen Sicherheitsforschungsprogramms KIRAS: Interimsevaluation 2011/2012. Wien/Berlin.

Schibany, A., Dinges, M., Reiner, C. (2013): Ex-post Evaluierung der Kompetenzzentrenprogramme Kplus und K_ind/K_net. Endbericht. Wien.

Schuch, K., Smoliner, S., Wagner, I., et al. (2013): Evaluierung der Forschungsk Kooperationen im Rahmen der bilateralen wissenschaftlich-technischen Abkommen und Vereinbarungen. Wien.

Stern, E. (2002) in: RTD Evaluation Toolbox - Assessing the Socio-Economic Impact of RTD-Policies

Terluin, I. J., Berkhout, P. (2011): The mixed case study approach. Assessing its usefulness as alternative to the CMEF. LEI Report 2011-038. The Hague

Warta, K., Geyer, A. (2011): Evaluierung des Humboldt-Forschungsstipendien-Programms der Alexander von Humboldt-Stiftung. Endbericht. Wien.

Warta, K., Geyer, A. (2011): Evaluation des Programms „COIN Cooperation & Innovation“. Endbericht. Wien.

Warta, K., Good, B., Geyer, A. (2009): Programmevaluierung Bridge. Endbericht. Wien.

Warta, K., Ohler, F., Konlechner, C., et al. (2014): Evaluierung des österreichischen Genomforschungsprogramms GEN-AU unter Einbe-

ziehung der Entwicklungen der Life Sciences Forschungslandschaft in Österreich. Endbericht. Wien.

Yin, Robert K. (2003): Case Study Research. Design and Methods. Third Edition. SAGE

AUTORINNEN

IRIS FISCHL

KMU Forschung Austria / Austrian Institute for SME Research
E: i.fischl@kmuforschung.ac.at

PETER KAUFMANN

KMU Forschung Austria / Austrian Institute for SME Research
E: p.kaufmann@kmuforschung.ac.at

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

KARIN GRASENICK, MAGDALENA PIERER, MARKUS GRUBER

UNTER WELCHEN RAHMENBEDINGUNGEN KÖNNEN EVALUIERUNGEN ZU VERÄNDERUNGEN BEITRAGEN?

Programme und Maßnahmen sind Ergebnisse eines komplexen Zusammenwirkens politischen Handelns. Sie entstehen in einem funktional differenzierten Netzwerk und werden in diesem Netzwerk umgesetzt. Evaluierungen (im Folgenden auch: EV) werden demnach durch unterschiedliche Institutionen beauftragt, die entsprechend ihrer Rolle innerhalb des Netzwerkes über bestimmte Möglichkeiten verfügen, deren Ergebnisse zu nutzen bzw. zu verwenden.

Im Zentrum dieses Artikels steht die Frage, unter welchen Bedingungen Evaluierungen Veränderungsprozesse unterstützen können, ohne per se dafür verantwortlich zu sein.

In der Literatur wird dazu primär auf die Qualität von Evaluierungen als wichtiger Erfolgsfaktor für deren Nutzung Bezug genommen. „Qualitätsvolle“ Evaluierungen sind notwendige, aber nicht hinreichende Voraussetzungen für die Akzeptanz und Nutzung für Veränderungen. Ebenso muss das institutionelle Umfeld und der Kontext, in dem Evaluierungen durchgeführt und Entscheidungen getroffen werden, beachtet werden. Damit Evaluierungen einen wichtigen Beitrag leisten können, müssen sie gezielt in Veränderungsprozesse eingebettet werden. Erfahrungen und Ansatzpunkte dazu bietet die Literatur zum Thema „Change Management“. Diese Literatur zeigt, welche Schritte für erfolgreiche Veränderungen notwendig sind und liefert wertvolle Hinweise, was Evaluierungen zu Veränderungsprozessen beitragen können.

Der folgende Artikel basiert auf einer interdisziplinären Auseinandersetzung mit Literatur und Reflexion von Erfahrungen, wobei die Themenbereiche „Evaluation“ und „Change Management in Organisationen“ kombiniert werden.

WAS IST EINE EVALUIERUNG?

Evaluierungen werden in diesem Artikel im Wesentlichen verstanden als die Bereitstellung von umfassenden Informationen und Entscheidungsgrundlagen für Stakeholder, damit Politik-Programme, Instrumente und Maßnahmen (rational) bewertet und weiterentwickelt werden können. Dieses Verständnis steht im Einklang mit Patton (1997), der (Programm-)Evaluierung definiert als *“the systematic collection of information about the activities, characteristics, and outcomes of programs to make judgments about the program, improve program effectiveness, and/or inform decisions about future programming”* (Patton 1997, S.23).

Auch Bortz & Döring (2003, S.102) betonen die *„systematische Anwendung empirischer Forschungsmethoden“*, um Programme, Instrumente und Maßnahmen in den unterschiedlichsten Phasen (Konzept, Implementierung, Wirksamkeit) zu bewerten. Die Überprüfung der Wirkung, der Effektivität und der Zielerreichung von Policies steht besonders bei Rossi et al. (2004) und Wollmann (2003), aber auch in der Definition der Europäischen Kommission (EU-Kommission 2000) im Vordergrund.

WARUM WERDEN EVALUIERUNGEN DURCHFÜHRT?

Evaluierung kann vielfältige Funktionen erfüllen: Erkenntnis und Information, Entwicklung und Lernen genauso wie Legitimation und Kontrolle (Wilhelm 2012; fteval 2012). Evaluierungen sollen helfen, *“(…) Handeln zu verstehen (Erkenntnisfunktion), Handeln zu verbessern (Entwicklungs- und Lernfunktion), Handeln darzustellen (Legitimationsfunktion) und Handeln zu überprüfen (Kontrollfunktion)”* (Wilhelm 2012, S.21).

Dabei wurde von einigen AutorInnen festgestellt, dass in der Praxis häufig die Legitimation im Vordergrund steht, also das Ablegen von Rechenschaft über den Einsatz von (knappen) Mitteln und die dadurch erzielten Resultate¹ (z.B. Mackay & Horton 2003). Evaluierungen haben jedoch auch eine wichtige Rolle als „learning tool“, also zum Lernen, Verbessern und Verändern von Politiken, Programmen und Maßnahmen (Measham 2009).

1 Auf die Problematik, ob und wie überhaupt Wirkungszusammenhänge hergestellt werden können (also inwiefern die beobachteten Resultate tatsächlich auf ein konkretes Projekt oder Programm zurückzuführen sind, und in welchem Ausmaß andere Einflussfaktoren eine Rolle spielen), kann hier nur am Rande hingewiesen werden. Es gibt in diesem Zusammenhang umfassende Diskussionen über Methoden und Vorgehensweisen: z.B. diskutiert Kromrey (2001) Voraussetzungen für eine „kausalanalytisch angeleitete Programmevaluations- und Wirkungsforschung“, und das Feldexperiment wird als „gold standard“ hervorgehoben (Rossi et al. 2004).

WIE WERDEN EVALUIERUNGEN GENUTZT?

In der Diskussion um die Verwendung und Sinnhaftigkeit von Evaluierung ist die „Nutzung“ (utilization, use) ein zentraler Aspekt (Fabian 2009; Wollmann 2013; Rossi et al. 2004)². Der Begriff der Nutzung bezieht sich auf die institutionelle Verankerung und den Stellenwert von Evaluierungen. Es geht darum, ob und wie Evaluierungen verwendet werden und einen Einfluss auf (politische) Entscheidungsprozesse haben, und durch welche Faktoren diese Verwendung beeinflusst wird (Schmidt 2009). In der Literatur werden unterschiedliche Ausprägungen der Nutzung unterschieden, die jedoch in der Realität typischerweise als „Mischformen“ vorkommen (nach Wollmann 2013; Fabian 2009; Schmidt 2009; Rossi et al. 2004):

- **Instrumentelle** Nutzung ist die direkte praktische, spezifische und dokumentierbare Verwendung von Evaluierungsergebnissen und Empfehlungen in Entscheidungsprozessen, zur Programmgestaltung oder zum Lösen von Problemen.
- **Prozessbezogene** Nutzung betont den Prozess des wechselseitigen Lernens und der Veränderung bereits während der Durchführung der Evaluierung, gefördert durch die Kommunikation und den Austausch zwischen Evaluierenden und Stakeholdern.
- Bei der **konzeptuellen** Nutzung löst das Wissen aus Evaluierungen einen Lernprozess aus und beeinflusst indirekt die Überzeugungen, Meinungen und Einstellungen der Policymaker, aber ohne spezifischen und dokumentierten Anlassfall.
- **Symbolische, politische oder legitimatorische** Nutzung hingegen bezeichnet die Verwendung von EV-Ergebnissen, um bereits bestehende Präferenzen oder Handlungen zu legitimieren, zu rechtfertigen und zu verteidigen. **Argumentativ** werden Evaluierungen genutzt, wenn sie z.B. in Politikdebatten als Argumente eingesetzt werden.
- **„Taktisch“/Strategisch** können EV-Ergebnisse genutzt werden, um Handlungsvorteile (z.B. Zeitgewinne) in einem politischen Konflikt zu gewinnen, zu beschwichtigen oder Koalitionen aufzubauen.

Im Folgenden steht die instrumentelle Nutzung, als direkteste Form der Verwendung von Evaluierungen, im Vordergrund. Eine instrumentelle Nutzung von Evaluierungen führt zu Veränderungen, die nicht nur fachlich-inhaltliche Auswirkungen haben, sondern insbesondere auch die Zusammenarbeit und die Aufgabenverteilung zwischen Institutionen, Agenturen und einzelnen Personen betreffen.

WAS BEEINFLUSST DIE NUTZUNG VON EVALUIERUNGSERGEBNISSEN?

Eine hohe Qualität der Evaluierung ist wesentliche Grundvoraussetzung für die Akzeptanz der formulierten Ergebnisse. Dementsprechend wurden und werden umfassende Qualitätsstandards und Kriterien für

die professionelle Durchführung von Evaluierungen formuliert. Häufig steht die wissenschaftlich korrekte, technisch-methodische Vorgehensweise im Vordergrund, aber auch Aspekte wie Glaubwürdigkeit und Plausibilität, Unabhängigkeit, Nachvollziehbarkeit und die Qualität der Daten werden thematisiert (siehe z.B. Patton 1999; fteval 2012). Verstärkt wird auch auf Aspekte der Anwendung bzw. Nutzung von EV hingewiesen (z.B. JCSEE 2011, "Utility Standards").

Das durch die Evaluierung bereitgestellte Wissen, die Einschätzungen und gegebenenfalls die Empfehlungen sollen in weiterer Folge durch die Auftraggebenden und das Politiknetzwerk, in das sie eingebettet sind, genutzt werden: als Basis für individuelles Lernen, für strategische Überlegungen und letztendlich für eine evidenzbasierte Politikgestaltung („evidence-based policy making“). Daten, ausgeklügelte Erhebungsmethoden und eine breitere Informationsbasis allein reichen aber nicht aus, um bessere und effizientere Policies zu machen (Borrás 2011; Sanderson 2002).

Evaluierungen haben häufig nur einen geringen Einfluss und werden sehr selektiv verwendet (Weiss et al. 2008; Borrás 2011; Højlund 2014). Wollmann (2013) stellt fest, dass EV-Ergebnisse tendenziell eher im Bereich der Durchführung von Maßnahmen genutzt werden, und kaum für größere, politisch sensible oder kontroverse Entscheidungen, wie etwa die Verteilung von finanziellen Ressourcen. Auch Teirlinck et al. (2013) beobachten eine (zu) schwache Verbindung zwischen Evaluierungen und Entscheidungsprozessen. Und Højlund (2014) spricht von einem Paradoxon, da das Ziel von Evaluierungen im Prinzip das Verbessern von Policies ist, es in der Praxis aber kaum zu solchen Veränderungen kommt. Diese Beobachtungen beziehen sich in einem großen Maß auf die instrumentelle, direkte Nutzung von Evaluierungen, die hier im Fokus der Betrachtung steht. Das heißt allerdings nicht, dass die Evaluierung nicht auf eine andere Art genutzt wird bzw. eine andere als die Entwicklungs- und Lernfunktion erfüllt.

Es gibt daher zahlreiche Publikationen, welche Erfolgsfaktoren und Hindernisse für die Nutzung von Evaluierungen analysieren. Ausgangspunkte sind dabei in der Regel die Evaluierung selbst, Fragen zu gewählten Methoden oder zur Prozessgestaltung, bis zu den erforderlichen Kompetenzen der Evaluierenden. Højlund (2014) kritisiert, dass durch diesen Fokus auf die Evaluierung an sich der organisationale Kontext häufig außer Acht gelassen wird. Gerade dieser Kontext ist aber besonders relevant: externer Druck und die interne Neigung, Evaluierungen durchzuführen, bestimmen die Nutzung von EV grundlegend mit (Højlund 2014). Im Folgenden werden die (partizipative) Prozessgestaltung sowie die Rolle des organisationalen und politischen Kontexts genauer betrachtet.

Sind partizipative Evaluierungen nützlicher?

Ein zentraler Erfolgsfaktor für Evaluierungen und die Nutzung ihrer Ergebnisse ist die Partizipation, also die Einbindung von möglichst vielen Stakeholdern (nicht nur „NutzerInnen“, sondern v.a. auch Policy-Maker, Programm-Manager, RegierungsvertreterInnen, PolitikerInnen) während des gesamten EV-Prozesses (Teirlinck et al. 2013). Patton betont mit sei-

2 Daneben wird in der Literatur vereinzelt auch die Perspektive der „Nützlichkeit“ unterschieden, also die Gestaltung der Evaluierung in einer Art und Weise, die den Informationsbedürfnissen der Zielgruppe(n) entspricht (Spiel & Bergsmann 2009). Andere AutorInnen argumentieren, dass statt Nutzung der weiter gefasste Begriff „EV-Einfluss“ besser geeignet ist, um unterschiedlichste Mechanismen der Verwendung und auch unbeabsichtigte, indirekte Auswirkungen zu erfassen (Lehtonen 2005; Henry & Mark 2003). Dabei wird betont, dass Evaluierungen (fast) immer in irgendeiner Form verarbeitet werden und zu einem Lernprozess beitragen.

ner „utilization-focused evaluation“ besonders, dass bereits während des Evaluierungsprozesses sehr viel gelernt wird, und nicht erst nach Abschluss der EV (Patton 1997).

Die aktive, umfassende und möglichst frühzeitige Einbindung von Stakeholdern und die enge Zusammenarbeit zwischen Evaluierenden und Evaluierten können das Wissen, die Einstellungen und Fähigkeiten aller involvierten Personen dauerhaft beeinflussen und sich merkbar auf zukünftige Entscheidungen und Handlungen auswirken (Forss et al. 2002; Horton & Mackay 2003; Wollmann 2013). Dieser Prozessnutzen kann auf unterschiedliche Weise zum Tragen kommen (Forss et al. 2002): Lernen zu lernen, fachgerechte Netzwerke entwickeln, ein gemeinsames Verständnis entwickeln, das Projekt oder Programm stärken, die Moral heben. Auch wenn also zum Beispiel ein Evaluierungsbericht nicht unmittelbar genutzt wird, hat sich die Evaluierung vielleicht darauf ausgewirkt, wie das Management einer Organisation über zukünftige Optionen denkt.

Der Erfolg der Einbindung von Stakeholdern wurde auch in empirischen Studien belegt. Zum Beispiel führen die stärkere Einbeziehung von Maßnahmenverantwortlichen und die amtsinterne Durchführung von Evaluierungen eher dazu, dass die Ergebnisse tatsächlich verwendet werden (Balthasar 2007, zitiert in Fabian 2009). Lake (2011) kommt in ihrer Studie ebenfalls zu dem Schluss, dass Kommunikation und Kooperation wesentliche Faktoren für den Nutzen und die Nutzung von Evaluierungen sind. Durch ein verbessertes Verständnis der einzelnen beteiligten Gruppen und die Entwicklung einer gemeinsamen Vorstellung über die Ziele und die entsprechend gewünschten Handlungsstrategien könnte die Nutzung verstärkt werden (Lake 2011). Das Entwickeln von gemeinsamen Definitionen und gegenseitigem Verständnis ist die Grundvoraussetzung für gute Kommunikation und eine erfolgreiche Evaluierung (Patton 1997).

Im Sinne eines formativen Zugangs zur Evaluierung wird also die Notwendigkeit von interaktiven EV-Systemen betont, die kontinuierlich mit den Prozessen der Politikformulierung und –umsetzung interagieren. Statt des Strebens nach „objektiven“ Evaluierungen (unabhängige Evaluierende, die Beweise produzieren, aber keine Empfehlungen), wird eine stärkere Involvierung der Evaluierenden in Lernprozesse mit allen Stakeholdern gefordert (Kuhlmann 1999, zitiert in Molas-Gallart & Davies 2006). Damit werden Evaluierende zu „Facilitators“, die neben der unabhängigen fachlichen Analyse auch Empfehlungen geben, auf die Informationsbedürfnisse der wesentlichen Stakeholder eingehen und so Entscheidungen und Veränderungen erleichtern (Horton & Mackay 2003; Molas-Gallart & Davies 2006).

Es gibt also vielfältige Ansätze und Ideen, wie Evaluierungen sinnvoller gestaltet und besser genutzt werden können, um tatsächlich einen Einfluss zu haben. Der Fokus auf Anwendungsorientierung und die Verpflichtung zum Nutzen und zur Nutzung der EV-Ergebnisse, wie sie in der Evaluationsforschung betont werden (Wilhelm 2012; Lake 2011; Patton 1997) stehen aber in einem gewissen Widerspruch zu Beobachtungen in der Praxis. Insgesamt scheinen die in der Literatur genannten Aspekte und Erfolgsfaktoren nur bedingt hilfreich oder wirksam, um das vorhandene Potenzial von Evaluierungen voll auszuschöpfen (Spiel & Bergsmann 2009). Wie diskutiert, wird primär bei der Evaluierung selbst und bei den Evaluierenden angesetzt. Damit verändert sich die Funkti-

on der Evaluierung und, im Falle einer partizipativen Evaluierung, auch deutlich die Rolle der Evaluierenden, denen damit implizit eine gewisse Verantwortung für Veränderungsprozesse zugesprochen wird.

Unklar bleibt dabei die Rolle der Auftraggebenden sowie das Zusammenspiel der AkteurInnen. Die Grenzen eines Evaluierungsprozesses im Unterschied zu einem Ergebnisverwertungs- und Implementierungsprozess verschwimmen dadurch und können Verwirrung stiften.

Welche Rolle spielt der Kontext?

Eine Reihe von Faktoren aus dem institutionellen, politischen und gesellschaftlichen Umfeld beeinflusst die gezielte Nutzung von Evaluierungen. Das sind zum Beispiel der organisationale und politische Kontext (z.B. Aktualität gewisser Inhalte und Themen, politischer Möglichkeitsraum, bürokratische Erfordernisse), aber auch Intentionen der Auftraggebenden (reine Rechtfertigung vs. andere Funktionen der EV), normative Verpflichtungen, die Notwendigkeit legitime Interessen zu bedienen; Werte, Weltanschauungen und Ideologie als treibende Kräfte im Policy-Prozess; die Verfügbarkeit von Ressourcen; Lobbyisten und Interessensgruppen (Lake 2011; Davies 2012; Meadowcroft & Steurer 2013; Teirlinck et al. 2013; Wollmann 2013). In komplexen politischen Entscheidungsprozessen, in denen Interessen und Anforderungen verhandelt und ausbalanciert werden müssen, sind die Ergebnisse von Evaluierungen oft nur eines von vielen Argumenten.

Evaluierungen sollten daher nicht als isolierter und eigenständiger Prozess betrachtet werden, sondern eingebettet in den Kontext des gesamten Politikzyklus. Durch diesen wiederkehrenden Zyklus aus Agenda Setting, Politikformulierung, Entscheidungsfindung, Politikimplementierung und Monitoring & Evaluierung sind die Rahmenbedingungen für Lern- und Veränderungsprozesse gegeben (Hildén 2011). Damit Evaluierungen in diesem Rahmen hilfreich sind, müssen sie aber aktiv eingebunden und von Anfang an im Prozess mitgedacht und mitgeplant werden. Geschieht das nicht, z.B. indem die Evaluierungsergebnisse keine Rolle in den ersten Schritten des darauffolgenden Politikzyklus spielen, kommt es zu einem Bruch in diesem Zyklus (Teirlinck et al. 2013). Das deutet auf Defizite in der institutionellen Verankerung des Lernens aus Erfahrung („lesson learning“) und der Politik-Verbesserung hin (Teirlinck et al. 2013).

Politiklernen wird unter anderem verstanden als dauerhafte Verhaltensänderungen, die auf neuen Informationen (oder Erfahrungen) beruhen, oder auch als spezifischer Prozess, in dem Wissen für die Entwicklung und Umsetzung von Policies verwendet wird (Biegelbauer 2007; Borrás 2011; Blum & Schubert 2009). Allerdings wird kaum diskutiert, WIE denn das Politiklernen beeinflusst und gefördert werden kann. VertreterInnen des „evidence-based policy making“ gehen davon aus, dass mehr Wissen und Information, bessere Daten etc. helfen. Das ist zwar notwendig und hilfreich, aber offensichtlich nicht ausreichend (Borrás 2011). Ausgehend von der Idee, dass der Lernprozess abhängig ist von den Eigenschaften und Merkmalen der Organisationen (als „agents of learning“), spricht Borrás (2011) von „**organizational capacities**“. Anhand eines transnationalen Vergleichs führt sie unterschiedliche Lernerfolge unter ähnlichen Rahmenbedingungen auf die in den einzelnen Ländern vorhandenen organisationalen Kapazitäten zurück. Dabei unterscheidet sie zwischen administrativer, analytischer und reflexiver/institutioneller Kapazität. Weitgehend unbeantwortet bleibt aber auch hier die Frage, wie diese Kapazitäten gebildet und gestärkt werden können.

Ansätze aus dem Change Management können dazu einen wertvollen Beitrag liefern.

WAS IST CHANGE MANAGEMENT?

Unter „Change Management“ bzw. Veränderungsmanagement werden alle *„absichtlichen, gesteuerten, organisierten und kontrollierten Anstrengungen zur antizipativen und zielgerichteten Organisationsgestaltung mit dem Ziel der Effektivitäts- und Effizienzsteigerung“* (Vahs 2005, S.249) verstanden – also das Anregen, Planen, Leiten und Durchführen von Veränderungsprozessen in Unternehmen und Organisationen. Dabei können diese Veränderungen geplant oder emergent, radikal oder inkrementell verlaufen, ganze Organisationen oder nur Teile davon (Strategien, Strukturen, Prozessabläufe, Technologien etc.) betreffen³ (Rees & Hall 2013). In jedem Fall sind dabei aber die handelnden Personen von besonderer Bedeutung, d.h. motivationale und emotionale Faktoren spielen in der Gestaltung eines Veränderungsprozesses eine wesentliche Rolle (Doppler & Voigt 2012).

Aus der Organisationstheorie kommend, bezieht sich das Konzept des Change Managements ursprünglich auf Organisationen, d.h. auf soziale Einheiten, deren Grenzen durch die Mitgliedschaft der AkteureInnen klar bestimmt sind, deren Zweck und Ziele formuliert werden können und die über formale Strukturen, Prozesse und Ressourcen zur Verteilung von Aufgaben und Verantwortlichkeiten sowie zur Regelung der Kommunikation verfügen (Kieser & Kubicek 1992).

Damit es aber zu wirklichem Wandel kommt, müssen die Rahmenbedingungen „passen“. In einer einfachen Formel⁴ bringen Dannemiller-Tyson-Associates (2000) diese Voraussetzung zum Ausdruck: $D * V * F > R$. Dabei steht D für „dissatisfaction“ (Unzufriedenheit mit der aktuellen Situation), V für „vision“ (Vision einer positiven Weiterentwicklung), F für „first steps“ (erste Schritte in Richtung der Vision), und R für „resistance“ (Widerstand gegenüber dem Wandel, Kosten des Wandels). Nur wenn das Produkt aus Unzufriedenheit, Vision und bereits vorhandenen ersten Schritten größer ist als der Widerstand, kann es zu Veränderungen kommen. Das heißt auch, dass alle drei Aspekte in einem gewissen Maß vorhanden sein müssen – wenn einer sehr klein ist (oder gar überhaupt fehlt), wird auch das Produkt sehr klein (bzw. Null), und wird vermutlich vom vorhandenen Widerstand übertroffen.

Wie in diesen Überlegungen zu den passenden Rahmenbedingungen für erfolgreiche Veränderungsprozesse deutlich wird, spielen Widerstand und die Überwindung von Widerstand eine wesentliche Rolle im Change Management (Doppler et al. 2002).

WIE ERFOLGREICH SIND GEPLANTE VERÄNDERUNGSPROZESSE IN ORGANISATIONEN?

Veränderungen in Unternehmen und Organisationen umzusetzen ist eine herausfordernde Aufgabe, die nicht immer erfolgreich ist. Manch-

mal werden, trotz klarer Signale für eine notwendige Neuorientierung, Veränderungen nicht oder zu spät in Angriff genommen. Es wurde festgestellt, dass ca. 70% der Wirtschafts-Unternehmen erst auf Veränderungen im Umfeld reagieren, wenn es bereits zu Einbußen bei Umsatz und Gewinn kommt (Kraus & Haghani 2004, zitiert in Lauer 2010). Von jenen Unternehmen, die aktiv einen Veränderungsprozess initiieren und umsetzen, geben laut einer Studie des Hernstein-Instituts 38% an, dass der Change-Prozess gescheitert ist (Befragung von 1000 deutschsprachigen Unternehmen, zitiert in Lauer (2010, S.41). By (2005) berichtet sogar von einer Misserfolgsquote von 70% der durchgeführten Change Programme.

Als wichtigste Ursache für dieses Scheitern von geplanten Veränderungsprozessen wird der Widerstand der Mitarbeitenden (30%) genannt, gefolgt von mangelhafter Prozess-Steuerung (25%) und zu raschem Tempo der Veränderung (20%) (Hernstein-Studie, zitiert in Lauer 2010, S.41f).

Die Ablehnung einer nach „objektiven“ Kriterien sinnvollen Veränderung und der Widerstand in einem laufenden Veränderungsprozess können unterschiedliche Ursachen haben – auf individueller oder auf organisationaler Ebene. Im Folgenden werden einige dieser ursächlichen Faktoren skizziert, die auch für die Akzeptanz von Evaluierungsergebnissen und deren institutionelle Nutzung relevant sind.

Lauer (2010) hält fest, dass auf individueller Ebene die Vermeidung von **„kognitiver Dissonanz“** eine wichtige Rolle spielt. Bei diesem Konzept der kognitiven Psychologie wird davon ausgegangen, dass Widersprüche in der Denkwelt des Menschen sehr unangenehm sind und vermieden werden. Menschen bemühen sich also um möglichst harmonische, widerspruchsfreie Bewusstseinsinhalte (Kognitionen). Dementsprechend werden widersprüchliche (dissonante) Informationen entweder ignoriert und gar nicht wahrgenommen, oder abgewertet und infrage gestellt. Ergänzend werden bestätigende, „konsonante“ Informationen gesucht, z.B. durch Gespräche mit ExpertInnen, von denen klar ist, dass sie die eigene Position vertreten. Hingegen sind Evaluierungsergebnisse, die nicht zu den vorhandenen Denkmustern, Werten und Vorstellungen passen, dissonante Informationen.

Eine weitere individuelle Verhaltensweise ist das sogenannte **„Satisficing“**. Damit ist die Tendenz gemeint, sich für die erstbeste Alternative, die als „gut genug“ für einen bestimmten Zweck eingeschätzt wird, zu entscheiden, und nicht nach vollständiger Optimierung zu streben. Dieses Verhalten hat mit unserer beschränkten Wahrnehmungs- und Verarbeitungsfähigkeit zu tun, die den Menschen dazu bringt, allzu komplexes Denken zu vermeiden. In diesem Zusammenhang ist für Veränderungsprozesse wesentlich, dass Menschen auch erst dann überhaupt nach Alternativen zu suchen beginnen, wenn sie bereits sehr „unzufrieden“ sind (Lauer 2010).

Auf organisationaler Ebene spielen Formalisierung, Hierarchie und Anreizmechanismen eine wichtige Rolle: Zentralisation, ein hoher Grad an Bürokratie in formalisierten Prozessen sowie Status und Einkommensunterschiede erschweren die Initiierung und Umsetzung von Veränderung (Hage & Aiken 1970, zitiert in Lauer 2010). Gerade in derartigen

3 In der Literatur lassen sich unzählige Ansätze und Modelle finden, die „Change“ und „Change Management“ kategorisieren, analysieren und operationalisieren (Brisson-Banks 2010; Pryor et al. 2008; By 2005; Kezar 2001).

4 Dabei handelt es sich um die Weiterentwicklung bzw. Vereinfachung einer Idee, die ursprünglich aus den 1960er Jahren von David Gleicher stammt, und von Richard Beckhard vorangetrieben wurde (Wheatley et al. 2003).

Organisationen kann Veränderung nur von entsprechenden zentralen Führungspersonen initiiert werden.

WELCHE SCHRITTE SIND FÜR ERFOLGREICHE VERÄNDERUNGSPROZESSE ZU BEACHTEN?

Unabhängig vom Grad der zu bewältigenden Veränderung, der Größe oder den Charakteristika einer Organisation, wurden aus der Analyse zahlreicher Praxisbeispiele Erfolgsfaktoren definiert und allgemein gültige Prozessschritte für die Planung und Umsetzung von Veränderungen formuliert (für Übersichten siehe z.B. Brisson-Banks 2010; Pryor et al. 2008; By 2005).

Eine wesentliche Basis sind die acht Stufen des Change Managements nach John P. Kotter (Kotter 2007). Er betont besonders die Rolle einzelner Personen, die die Notwendigkeit sehen etwas zu ändern und die Initiative ergreifen („leadership“).

1. **Establishing a sense of urgency:** Veränderungsbedarf wird durch fachliche Analysen begründet. Diese Analysen müssen in ihrer Bedeutung erkannt und kommuniziert werden. Der Bedarf einer Veränderung muss einer möglichst großen Anzahl an Personen unbedingt erforderlich erscheinen. Ist dies nicht der Fall, fehlt die Motivation, sich an einer Umsetzung zu beteiligen und die damit verbundenen Risiken und Unsicherheiten zu überwinden.
2. **Forming a powerful guiding coalition:** Ein Veränderungsprozess bedarf der Zustimmung – und der aktiven Beteiligung jener Personen, die über die hierfür notwendigen Entscheidungsbefugnisse verfügen und entsprechende Ressourcen für die Umsetzung sicherstellen können. Diese Personen müssen den Veränderungsprozess verantworten und ihn während der gesamten Laufzeit als Führungsteam begleiten.
3. **Creating a vision:** Eine angestrebte Veränderung bedarf einer Vision, die einfach und verständlich vermittelt, was durch die Veränderung erreicht werden kann. Dieses „Bild mit Zugkraft“ ist knapp formuliert und bildet die Basis für die Kommunikation und die Einbindung aller relevanten Stakeholder.
4. **Communicating the vision:** Das Führungsteam muss alle vorhandenen Mittel und Möglichkeiten nutzen, um Vision und Strategie zu kommunizieren. Das Handeln des Führungsteams muss diesem „Bild mit Zugkraft“ entsprechen.
5. **Empowering others to act on the vision:** Veränderungsprozesse werden am besten durch eine breite Beteiligung all jener Personen gestaltet, welche die neuen Prozesse in weiterer Folge tatsächlich umsetzen werden. Es gilt, Hindernisse aus dem Weg zu räumen, die Risikobereitschaft zu fördern und zu neuen Ideen, Aktivitäten und Aktionen zu ermutigen.
6. **Planning for and creating short-term wins:** Verbesserungen, die kurzfristig erreicht werden können, fördern die Motivation. Das Engagement von Mitarbeitenden für diese sichtbaren Erfolge muss anerkannt und belohnt werden.
7. **Consolidating improvements and producing still more change:** Kleine Erfolge machen langfristige Ziele glaubwürdig. Es gilt, die entstandene Motivation zu nutzen, um Systeme, Strukturen und Policies zu ändern, die nicht zur Vision passen. Personalentwicklung und Training treiben die Umsetzung weiter voran. Neue Teil-Projekte, Themen und Change Agents beleben den Prozess wieder.
8. **Institutionalizing new approaches:** Bisherige, aus Erfahrung entstandene Verhaltensweisen sind sehr beständig. Es ist daher wichtig, den Zusammenhang zwischen den neuen Verhaltensweisen und dem (Unternehmens-)erfolg zu betonen. Das neue Wissen, die neuen Verhaltensweisen müssen vor allem auch von allen bestehenden und zukünftigen Führungskräften umgesetzt werden.

WAS KÖNNEN EVALUIERUNGEN FÜR VERÄNDERUNGSPROZESSE „TUN“?

Abhängig vom Gegenstand, dem Zeitpunkt, den Inhalten, bzw. allgemein dem Ziel und Zweck einer Evaluierung gibt es unterschiedliche Vorstellungen darüber, was eine Evaluierung leisten soll und welche Aufgaben und Zuständigkeiten in den Verantwortungsbereich der Evaluierenden fallen. Wenn eine direkte, instrumentelle Nutzung der Evaluierung und ihrer Ergebnisse angestrebt wird, kann die Evaluierung dazu beitragen, den dafür notwendigen Veränderungsprozess auszulösen. Evaluierungen bzw. Evaluierungsergebnisse und –empfehlungen sind also wichtige Voraussetzung und Ausgangspunkt für Veränderungen. Die konkrete Planung und Umsetzung der Veränderungsprozesse ist jedoch nicht (mehr) Aufgabe der Evaluierung.

- Evaluierungen stellen umfassende Informationen, Daten und Empfehlungen bereit, die als Basis für Entscheidungen dienen. Die Qualität der Evaluierung und die Qualifikation der Evaluierenden können zu einer erfolgreichen Nutzung beitragen, da ein nachvollziehbarer Prozess der Ergebnisfindung die Akzeptanz von Entscheidungsgrundlagen erhöht (Rossi et al. 2004).
- Die Gestaltung des EV-Prozesses bietet Potenzial, die Betroffenen auf mögliche Veränderungen vorzubereiten bzw. gemeinsam mit ihnen darauf hinzuwirken: Partizipative Evaluierungen können durch die breite Einbindung von Stakeholdern und die damit einhergehende Kommunikation einen wichtigen Beitrag zur Entwicklung einer geeigneten Ausgangsbasis für die notwendige Kommunikation und Zusammenarbeit im späteren Veränderungsprozess schaffen.
- Nach Kotters Ansatz können Evaluierungen hilfreich sein, um einen „sense of urgency“ zu vermitteln. Das Entwickeln eines gemeinsamen Verständnisses und eines strategischen Fokus ist wichtig, um dann in weiterer Folge zu überlegen, welche Bedeutung Daten und Erkenntnisse haben, und welche Möglichkeiten es zur weiteren Entwicklung gibt (Dannemiller-Tyson-Associates 2000).
- Werden in einer Evaluierung Empfehlungen formuliert, so können diese als Ausgangspunkt herangezogen werden, um eine

Vision oder erste konkrete Schritte sichtbar zu machen, welche die handelnden Personen motivieren, neue Ideen zu entwickeln.

Insgesamt bedeutet das für die Evaluierenden in erster Linie, ein Bewusstsein auszubilden und „Change-Aspekte“ soweit als möglich bei der Definition der zu erwartenden Ergebnisse und dem Evaluierungsdesign zu berücksichtigen. In diesem Kontext ist es sinnvoll, teils gemeinsam mit den Auftraggebenden, vor der Evaluierung einige Fragen zu klären, mit deren Hilfe die Chancen auf Realisierung abgeschätzt werden können:

1. Welche Veränderung wird angestrebt? Handelt es sich um inkrementelle Verbesserungen oder um den Wunsch einer radikalen Neuplanung?
2. Welche Institutionen, Politikfelder und Personen müssen für eine erfolgreiche Umsetzung beteiligt sein? Lassen sich Interessen, der zu erwartende Nutzen und die befürchteten Verluste der Beteiligten abschätzen?
3. Wer sollte die Evaluierung beauftragen, um die Akzeptanz der Ergebnisse sicherzustellen? Verfügen die Personen, welche die Evaluierung beauftragen, über die notwendigen Entscheidungs- und Handlungsspielräume, um den daraus resultierenden Veränderungs-Prozess tatsächlich voranzutreiben?
4. Wie sind die Voraussetzungen für gelungene Veränderungen aus der Perspektive der einzubindenden Personen und Institutionen zu beurteilen?
5. Welche Rolle kann die Evaluierung in dem zu planenden Veränderungsprozess spielen?

WELCHE HERAUSFORDERUNGEN ERGEBEN SICH FÜR DIE EINBETTUNG VON EVALUIERUNGEN IN VERÄNDERUNGSPROZESSE?

Für die erfolgreiche instrumentelle Nutzung von Evaluierungen bedarf es der Planung und Umsetzung geeigneter Veränderungsprozesse. Kenntnis und Kompetenzen im Bereich des Change Management können dazu beitragen, dass Evaluierungen ihre grundlegendsten Funktionen (Lernen und Entwickeln) besser erfüllen können.

Evaluierungen können „veränderungsfördernd“ sein, wenn sie gezielt und durchdacht in einen übergeordneten Policy-Making und Veränderungsprozess eingebettet sind. Für die Forschungs- und Technologiepolitik bedeutet dies, die Evaluierung im gesamten Politikzyklus mitzudenken und bewusst zu planen.

Die Gestaltung des Veränderungsprozesses ist jedoch nicht Aufgabe der Evaluierung. Sie liegt in der Hand von Stakeholdern (Auftraggebende und EntscheidungsträgerInnen), die über die notwendigen Handlungsspielräume im Politikzyklus verfügen und als treibende Kräfte verantwor-

tung übernehmen. Die Veränderungsbereitschaft dieser Personen ist die Voraussetzung für die tatsächliche Wirkung von Evaluierungen.

Dazu stellen sich folgende Fragen für eine weiterführende Diskussion:

- Braucht es nicht viel stärker eine aktive Gestaltung von zwei Prozesszyklen, deren Ineinandergreifen verstärkt in der Konzeption Beachtung finden sollte? Also die Gestaltung des Politikzyklus und der dort durch EntscheidungsträgerInnen vorzunehmenden Veränderungsprozesse einerseits und der Evaluierung andererseits?
- Können Prozesse ohne die Rolle eines aktiven „Change Agents“ zu erfolgreichen Veränderungen führen? Was kann aus der Change Management Erfahrung auch für Politikprozesse gelernt werden?
- Ist die zunehmende Anforderung an Evaluierende, als „Facilitators“ oder sogar als „Change Agents“ zu agieren, nicht eine deutliche Überforderung der Erwartungen und Möglichkeiten?
- Werden in den „Terms of Reference“ immer relevante Fragestellungen im Hinblick auf mögliche Veränderungen gestellt?

Die instrumentelle Umsetzung von Empfehlungen einer Evaluierung kann jedoch die Mitwirkung von AkteurInnen aus unterschiedlichen Institutionen und Politikbereichen erfordern. Die AkteurInnen sind dabei in geringerem Maße auf ein gemeinsames Ziel ausgerichtet als dies innerhalb ein und derselben Organisation möglich ist. Dennoch ist es Aufgabe des Veränderungsprozesses, durch eine möglichst breite Einbindung aller beteiligten Personen und ihrer Expertise, im Detail die passenden Lösungen für eine gelungene Implementierung zu finden. Im Vergleich zu geplanten Veränderungen in Organisationen vergrößern sich so die Herausforderungen in komplexen FTI-politikrelevanten Netzwerken mit ihren vielfältigen, funktional differenzierten Ebenen und AkteurInnen. Damit verliert die bewusste Gestaltung von Veränderungen jedoch nicht an Bedeutung. Die Erkenntnisse aus dem Change Management sind hier sogar besonders hilfreich, um Widerstände besser zu verstehen und die notwendigen Voraussetzungen für eine gelungene Umsetzung von Empfehlungen zu schaffen.

LITERATURVERZEICHNIS

- Balthasar, A.**, 2007. Institutionelle Verankerung und Verwendung von Evaluationen. Praxis und Verwendung von Evaluationen in der schweizerischen Bundesverwaltung., Zürich: Rüegger.
- Biegelbauer, P.**, 2007. Ein neuer Blick auf politisches Handeln : Politik-Lernansätze im Vergleich. Österreichische Zeitschrift für Politikwissenschaft (ÖZP), 36(3), pp.231–247.
- Blum, S. & Schubert, K.**, 2009. Politikfeldanalyse, Wiesbaden: VS Verlag für Sozialwissenschaften.
- Borrás, S.**, 2011. Policy learning and organizational capacities in innovation policies. Science and Public Policy, 38(9), pp.725–734.
- Bortz, J. & Döring, N.**, 2003. Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler, Berlin/Heidelberg/New York: Springer Verlag.
- Brisson-Banks, C. V.**, 2010. Managing change and transitions: a comparison of different models and their commonalities. Library Management, 31(4/5), pp.241–252.
- By, R.T.**, 2005. Organisational change management: A critical review. Journal of Change Management, 5(4), pp.369–380.
- Dannemiller-Tyson-Associates**, 2000. Whole-scale change. Unleashing the magic in organizations, San Francisco: Berrett-Koehler Publishers, Inc.
- Davies, P.**, 2012. The State of Evidence-Based Policy Evaluation and its Role in Policy Formation. National Institute Economic Review, 219(1), pp.R41–R52.
- Doppler, K. et al.**, 2002. Unternehmenswandel gegen Widerstände. Change Management mit den Menschen, Frankfurt am Main: campus Verlag.
- Doppler, K. & Voigt, B.**, 2012. Feel the Change! Wie erfolgreiche Change Manager Emotionen steuern, Frankfurt am Main: campus Verlag.
- EU-Kommission**, 2000. Reforming the Commission - A White Paper - Part I, Available at: [http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:52000DC0200\(01\):EN:HTML](http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:52000DC0200(01):EN:HTML).
- Fabian, C.**, 2009. Nutzung der Evaluationsfunktion im Vergleich. In T. Widmer, W. Beywl, & C. Fabian, eds. Evaluation. Ein systematisches Handbuch. Wiesbaden: VS Verlag für Sozialwissenschaften, pp. 498–505.
- Forss, K., Rebien, C.C. & Carlsson, J.**, 2002. Process Use of Evaluations: Types of Use that Precede Lessons Learned and Feedback. Evaluation, 8(1), pp.29–45.
- fteval**, 2012. Standards der Evaluierung in der Forschungs- und Technologiepolitik. Available at: http://www.fteval.at/upload/fteval_Standards_deutsch.pdf.
- Hage, J. & Aiken, M.**, 1970. Social Change in Complex Organizations, New York: Random House.
- Henry, G.T. & Mark, M.M.**, 2003. Beyond use: Understanding evaluation's influence on attitudes and actions. American Journal of Evaluation, 24(3), pp.293–314.
- Hildén, M.**, 2011. The evolution of climate policies – the role of learning and evaluations. Journal of Cleaner Production, 19(16), pp.1798–1811.
- Højlund, S.**, 2014. Evaluation use in the organizational context - changing focus to improve theory. Evaluation, 20(1), pp.26–43.
- Horton, D. & Mackay, R.**, 2003. Using evaluation to enhance institutional learning and change: recent experiences with agricultural research and development. Agricultural Systems, 78(2), pp.127–142.
- JCSEE**, 2011. Program Evaluation Standards Statements, Available at: <http://www.jcsee.org/program-evaluation-standards-statements> [Accessed September 5, 2014].
- Kezar, A.J.**, 2001. Understanding and Facilitating Organizational Change in the 21st Century: Recent Research and Conceptualizations, San Francisco: Jossey-Bass.
- Kieser, A. & Kubicek, H.**, 1992. Organisation 3. ed., Berlin: de Gruyter.
- Kotter, J.P.**, 2007. Leading Change. Why transformation efforts fail. Harvard Business Review, (January), pp.96–103.
- Kraus, K.-J. & Haghani, S.**, 2004. Krisenverlauf und Krisenbewältigung – der aktuelle Stand. In N. et al Bickhoff, ed. Die Untern ehmenkrise als Chance. Innovative Ansätze zur Sanierung und Restrukturierung.
- Kromrey, H.**, 2001. Evaluation – ein vielschichtiges Konzept. Begriff und Methodik von Evaluierung und Evaluationsforschung. Empfehlungen für die Praxis. Sozialwissenschaften und Berufspraxis, 24(2), pp.105–131.
- Kuhlmann, S.**, 1999. Distributed intelligence: Combining evaluation, foresight and technology assessment. Institute for Prospective and Technological Studies Report, 40, pp.16–22.
- Lake, A.**, 2011. Rationality and Utility of Evaluation Research in Policy Making. Studies of Transition States and Societies, 3(3), pp.64–82.
- Lauer, T.**, 2010. Change Management. Grundlagen und Erfolgsfaktoren, Heidelberg/Dordrecht/London/New York: Springer Verlag.
- Lehtonen, M.**, 2005. Environmental Policy Evaluation in the Service of Sustainable Development : Influence of the OECD Environmental Performance Reviews from the Perspective of Institutional Economics. Université de Versailles Saint-Quentin-en-Yvelines.
- Mackay, R. & Horton, D.**, 2003. Expanding the use of impact assessment and evaluation in agricultural research and development. Agricultural Systems, 78(2), pp.143–165.

Meadowcroft, J. & Steurer, R., 2013. Assessment Practices in the Policy and Politics Cycles: A Contribution to Reflexive Governance for Sustainable Development? *Journal of Environmental Policy & Planning*, (September 2014), pp.1–23.

Measham, T.G., 2009. Social learning through evaluation: a case study of overcoming constraints for management of dryland salinity. *Environmental management*, 43(6), pp.1096–1107.

Molas-Gallart, J. & Davies, A., 2006. Toward Theory-Led Evaluation: The Experience of European Science, Technology, and Innovation Policies. *American Journal of Evaluation*, 27(1), pp.64–82.

Patton, M.Q., 1999. Enhancing the quality and credibility of qualitative analysis. *Health services research*, 34(5 Pt 2), pp.1189–1208.

Patton, M.Q., 1997. *Utilization-focused Evaluation*. The New Century Text 3rd ed., Thousand Oaks / London / New Delhi: SAGE Publications.

Pryor, M.G. et al., 2008. Challenges facing change management theories and research. *Delhi Business Review*, 9(1).

Rees, G. & Hall, D., 2013. *Managing Change*. In *Leading, Managing and Developing People*.

Rossi, P.H., Lipsey, M.W. & Freeman, H.E., 2004. *Evaluation. A Systematic Approach* 7th ed., Thousand Oaks / London / New Delhi: SAGE Publications.

Sanderson, I., 2002. Evaluation, policy learning and evidence-based policy making. , 80(1), pp.1–22.

Schmidt, C., 2009. Institutionalisation von Evaluation innerhalb der EU-Politik zur Entwicklung des ländlichen Raums: Umsetzung und Verwendung im politischen Entscheidungsprozess. Universität Rostock. Available at: http://rosdok.uni-rostock.de/file/rosdok_disshab_0000000418/rosdok_derivate_0000004252/Dissertation_Schmidt_2010.pdf.

Spiel, C. & Bergmann, E., 2009. Zur Nutzung der Evaluationsfunktion: Partielle Bestandsaufnahme und Pilotstudie in österreichischen und deutschen Bundesministerien. In T. Widmer, W. Beywl, & C. Fabian, eds. *Evaluation. Ein systematisches Handbuch*. Wiesbaden: VS Verlag für Sozialwissenschaften, pp. 477–485.

Teirlinck, P. et al., 2013. Closing the policy cycle: Increasing the utilization of evaluation findings in research, technological development and innovation policy design. *Science and Public Policy*, 40(3), pp.366–377.

Vahs, D., 2005. *Organisation. Einführung in die Organisationstheorie und -praxis*. Stuttgart: Schäffer-Poeschel Verlag.

Weiss, C.H. et al., 2008. The Fairy Godmother--and Her Warts: Making the Dream of Evidence-Based Policy Come True. *American Journal of Evaluation*, 29(1), pp.29–47.

Wheatley, M. et al., 2003. *Organization Development at Work. Conversations on the Values, Applications, and Future of OD*, San Francisco:

Preiffer. Available at: http://www.untag-smd.ac.id/files/Perpustakaan_Digital_2/ORGANIZATIONAL_CHANGE_Organization_development_at_work_conversations_on_the_values_applications.pdf.

Wilhelm, J.L., 2012. *Wozu Evaluation? Organisationssysteme bewerten Stadtteilfeörderung mit Kalkül*, Potsdam: Universitätsverlag Potsdam.

Wollmann, H., 2013. Die Untersuchung der (Nicht-)Verwendung von Evaluationsergebnissen in Politik und Verwaltung. Ein vernachlässigte Fragestellung der Evaluationforschung. *der moderne staat - Zeitschrift für Public Policy, Recht und Management*, Sonderheft, pp.87–103.

Wollmann, H., 2003. Kontrolle in Politik und Verwaltung: Evaluation, Controlling und Wissensnutzung. In N. C. Schubert, K.; Bandelow, ed. *Lehrbuch der Politikfeldanalyse*. München: Verlag Oldenbourg, pp. 335–360.

AUTORINNEN

KARIN GRASENICK

convelop cooperative knowledge design gmbh.

E: karin.grasenick@convelop.at

MAGDALENA PIERER

convelop cooperative knowledge design gmbh.

E: magdalena.pierer@convelop.at

MARKUS GRUBER

convelop cooperative knowledge design gmbh.

E: markus.gruber@convelop.at

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 Brey

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



ISSUE 38

PUBLIC PROCUREMENT IN SEE INNOVATION EVALUATIONS: A COMPARATIVE AND NEEDS ASSESSMENT STUDY
Lena Tspouri, Nikos Sidiropoulos

HOW TO DESIGN THE CONCEPTUAL LINK BETWEEN BIBLIOMETRIC ARTICLE ANALYSIS AND R&D STATISTICS BASED ON OECD FRASCATI MANUAL
David F. J. Campbell

SELF-ASSESSMENT METHODOLOGY FOR E-INFRASTRUCTURES' IMPACT ASSESSMENT CHALLENGES, RESULTS AND RECOMMENDATIONS OF ERINA+
Josef Benedikt

RESULTS FROM AN INTERNATIONAL BENCHMARKING EXERCISE OF THE STRATEGIC CENTRES FOR SCIENCE, TECHNOLOGY AND INNOVATION IN FINLAND
Michael Dinges

ZWISCHENEVALUIERUNG DES STRATEGIEPROGRAMMS FÜR MOBILITÄTS- UND VERKEHRSTECHNOLOGIE (IV2SPPLUS) - EINE ANALYSE DER

TEILNEHMER UND AUSGEWÄHLTER ASPEKTE DES TEILNEHMERVERTHALTENS
Iris Fischl, Thomas Oberholzner

BUILDING BRIDGES, TYING NETS: HOW TO PROMOTE RESEARCH COOPERATIONS? RESULTS FROM THE EVALUATIONS OF BRIDGE AND COIN
Katharina Warta

ZUR NEUORDNUNG DER EXTERNEN QUALITÄTSSICHERUNG IM ÖSTERREICHISCHEN HOCHSCHULSYSTEM
Alexander Kohler

EIN RESÜMEE ZUR PLATTFORMVERANSTALTUNG - „ZIELE UND BEWERTUNGSKRITERIEN FTI-POLITISCHER PROGRAMME“
Thomas Jud

BUCHBESPRECHUNG VON: CAMPBELL, DAVID F. J. / ELIAS G. CARAYANNIS (2013). EPISTEMIC GOVERNANCE IN HIGHER EDUCATION. QUALITY ENHANCEMENT OF UNIVERSITIES FOR DEVELOPMENT
Thorsten D. Barth



ISSUE 37

EVERYTHING YOU ALWAYS WANTED TO KNOW ABOUT EVALUATION, BUT WERE NEVER TOLD
Karin Grasenick, Stephan Kupsa

THE LEGITIMACY OF A FUNDING AGENCY IN BASIC RESEARCH. SOME EMPIRICAL EVIDENCES FROM THE AUSTRIAN SCIENCE FUND (FWF)
Christian Fischer, Falk Reckling

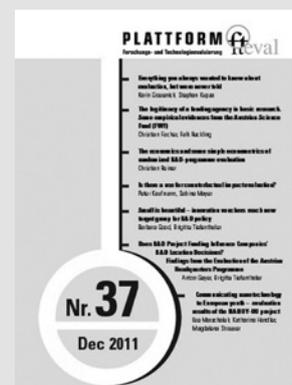
THE ECONOMICS AND SOME SIMPLE ECONOMETRICS OF RANDOMIZED R&D-PROGRAMME EVALUATION
Christian Reiner

IS THERE A USE FOR COUNTERFACTUAL IMPACT EVALUATION?
Peter Kaufmann, Sabine Mayer

SMALL IS BEAUTIFUL – INNOVATION VOUCHERS REACH NEW TARGET GROUP FOR R&D POLICY
Barbara Good, Brigitte Tiefenthaler

DOES R&D PROJECT FUNDING INFLUENCE COMPANIES' R&D LOCATION DECISIONS? FINDINGS FROM THE EVALUATION OF THE AUSTRIAN HEADQUARTERS PROGRAMME
Anton Geyer, Brigitte Tiefenthaler

COMMUNICATING NANOTECHNOLOGY TO EUROPEAN YOUTH – EVALUATION RESULTS OF THE NANOY-OU PROJECT
Ilse Marschalek, Katharina Handler, Magdalena Strasser



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, Technopolis Austria, Vienna Science and Technology Fund (WWTF), The Technology Agency of the City of Vienna (ZIT), 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 2014