

50

fteval JOURNAL for Research and Technology Policy Evaluation



ISSUE 50 | MARCH 2020

ISSUE №50

JOHANNES GADNER AND
JÜRGEN JANGER

MONITORING NATIONAL
INNOVATION PERFORMANCE:
GOING BEYOND
STANDARDIZED RANKINGS

MICHAEL DINGES, ANNA WANG
AND KLAUS SCHUCH

USING THE DELPHI METHOD
IN EVALUATIONS –
INCORPORATING A FUTURE
ORIENTED PERSPECTIVE IN
EVALUATIONS

RICHARD HEIDLER

FUNDING RESEARCH DATA
INFRASTRUCTURES:
FUNDING CRITERIA IN
GRANT PEER REVIEW

FEDERAL MINISTRY OF EDUCATION,
SCIENCE AND RESEARCH
Minoritenplatz 5, 1014 Vienna
Mag.^a Irene Danler
E: irene.danler@bmbwf.gv.at
Mag.^a Simone Mesner
E: simone.mesner@bmbwf.gv.at

FEDERAL MINISTRY FOR DIGITAL
AND ECONOMIC AFFAIRS
Stubenring 1, 1014 Vienna
Mag.^a Sabine Pohoryles-Drexel
E: sabine.pohoryles-drexel@bmfwf.gv.at

FEDERAL MINISTRY FOR CLIMATE
ACTION, ENVIRONMENT, ENERGY,
MOBILITY, INNOVATION AND
TECHNOLOGY
Radetzkystraße 2, 1030 Vienna
Dr. Rupert Pichler
E: rupert.pichler@bmk.gv.at
Dr. Mario Steyer
E: mario.steyer@bmk.gv.at

ACR – AUSTRIAN COOPERATIVE
RESEARCH
Sensengasse 1, 1010 Vienna
Dr.ⁱⁿ Sonja Sheikh
E: sheikh@acr.ac.at

AUSTRIAN COUNCIL FOR RESEARCH
AND TECHNOLOGY DEVELOPMENT
Pestalozzigasse 4/DG 1, 1010 Vienna
Dr. Johannes Gadner
E: j.gadner@rat-fte.at

AIT – AUSTRIAN INSTITUTE OF
TECHNOLOGY
Giefinggasse 4, 1210 Vienna
Mag. Michael Dinges
E: michael.dinges@ait.ac.at
Mag.^a Barbara Heller-Schuh, MAS
E: barbara.heller-schuh@ait.ac.at

AQ AUSTRIA – AGENCY FOR QUALITY
ASSURANCE AND ACCREDITATION
AUSTRIA
Franz-Klein-Gasse 5, 1190 Vienna
Dr.ⁱⁿ Elisabeth Froschauer-Neuhauser
E: elisabeth.froschauer-neuhauser@aq.ac.at
Dr.ⁱⁿ Eva Maria Freiberger
E: eva.maria.freiberger@aq.ac.at

AWS – AUSTRIA
WIRTSCHAFTSSERVICE GMBH
Walcherstraße 11A, 1020 Vienna
Mag.^a Marlis Baurecht
E: M.Baurecht@aws.at
Mag. Norbert Knoll
E: n.knoll@aws.g.at

CDG – CHRISTIAN DOPPLER
RESEARCH ASSOCIATION
Boltzmannngasse 20, 1090 Vienna
DIⁱⁿ Mag.^a Brigitte Müller
E: brigitte.mueller@cdg.ac.at

CONVELOP – COOPERATIVE
KNOWLEDGE DESIGN GMBH
Bürgergasse 8-10/1, 8010 Graz
DIⁱⁿ Dr.ⁱⁿ Karin Grasenick
E: karin.grasenick@convelop.at
Erdbergstraße 82/4, 1030 Wien
Mag. Thomas Jud
E: thomas.jud@convelop.at

FFG – AUSTRIAN RESEARCH
PROMOTION AGENCY
Haus der Forschung,
Sensengasse 1, 1090 Vienna
DIⁱⁿ Dr.ⁱⁿ Sabine Mayer
E: sabine.mayer@ffg.at
Mag. Leonhard Jörg
E: leonhard.joerg@ffg.at

FWF – AUSTRIAN SCIENCE FUND
Haus der Forschung
Sensengasse 1, 1090 Vienna
Dr. Falk Reckling
E: falk.reckling@fwf.ac.at
Dr. Thomas Völker
E: thomas.voelker@fwf.ac.at

IHS – INSTITUTE FOR ADVANCED
STUDIES
Josefstädter Straße 39, 1080 Vienna
Dr.ⁱⁿ Angela Wroblewski
E: wroblews@ihs.ac.at
Mag. Richard Sellner
E: richard.sellner@ihs.ac.at

INDUSTRIEWISSENSCHAFTLICHES
INSTITUT – IWI
Mittersteig 10, 1050 Wien
FH-Hon.Prof. Dr. Dr. Herwig W. Schneider
E: schneider@iwi.ac.at
Mag. Philipp Brunner
E: brunner@iwi.ac.at

JOANNEUM RESEARCH
FORSCHUNGSGESELLSCHAFT MBH
Haus der Forschung
Sensengasse 1, 1090 Vienna
Mag. Wolfgang Polt
E: wolfgang.polt@joanneum.at
Mag. Jürgen Streicher
E: juergen.streicher@joanneum.at

KMU FORSCHUNG AUSTRIA –
AUSTRIAN INSTITUTE FOR SME
RESEARCH
Gusshausstraße 8, 1040 Vienna
Mag.^a Iris Fischl
E: i.fischl@kmuforschung.ac.at
Mag. Peter Kaufmann
E: p.kaufmann@kmuforschung.ac.at

LUDWIG BOLTZMANN GESELLSCHAFT
Nußdorfer Straße 64, 1090 Vienna
Mag. Patrick Lehner
E: patrick.lehner@lbg.ac.at

OEAW – AUSTRIAN ACADEMY OF
SCIENCE
Dr. Ignaz Seipel-Platz 2, 1010 Vienna
Nikolaus Göth, MSc
E: nikolaus.goeth@oeaw.ac.at

TECHNOPOLIS GROUP
AUSTRIA
Rudolfsplatz 12/11, 1010 Vienna
Mag.^a Katharina Warta
E: warta@technopolis-group.com
Tobias Dudenbostel, MA
E: tobias.dudenbostel@technopolis-group.com

VIENNA BUSINESS AGENCY.
A SERVICE OFFERED BY THE CITY OF
VIENNA.
Mariahilfer Straße 20, 1070 Vienna
Robert Mayer-Unterholzer
E: mayer-unterholzer@wirtschaftsagentur.at

WIFO – AUSTRIAN INSTITUTE OF
ECONOMIC RESEARCH
Arsenal, Objekt 20
Postfach 91, 1103 Vienna
Dr. Jürgen Janger
E: juergen.janger@wifo.ac.at
Dr. Andreas Reinstaller
E: andreas.reinstaller@wifo.ac.at

WPZ RESEARCH GMBH
Mariahilfer Straße 115/16, 1060 Vienna
Dr.ⁱⁿ Brigitte Ecker
E: brigitte.ecker@wpz-research.com

WWTF – VIENNA SCIENCE AND
TECHNOLOGY FUND
Schlickgasse 3/12, 1090 Vienna
Dr. Michael Stampfer
E: michael.stampfer@wwtf.at
Dr. Michael Strassnig
E: michael.strassnig@wwtf.at

ZSI – CENTRE FOR SOCIAL INNOVATION
Linke Wienzeile 246, 1150 Vienna
Dr. Klaus Schuch
E: schuch@zsi.at
Dr. Dorothea Sturn
E: sturn@zsi.at

INHALTSVERZEICHNIS

- 3 EDITORIAL
- 4 MONITORING NATIONAL INNOVATION PERFORMANCE:
GOING BEYOND STANDARDIZED RANKINGS
JOHANNES GADNER AND JÜRGEN JANGER
- 11 ZUGANG ZU REGISTER- UND INDIVIDUALDATEN FÜR DIE
WISSENSCHAFTLICHE FORSCHUNG IN ÖSTERREICH
GERHARD SCHWARZ, THOMAS KÖNIG, HARALD OBERHOFER UND MICHAEL STRASSNIG
- 16 NEUES WAGEN – MIT DER SANDPIT-METHODE ZU NEUEN
KONSTELLATIONEN & INTERDISZIPLINÄREN IDEEN
SANDPIT-ERFAHRUNGEN AUS SICHT EINER FÖRDERUNGSAGENTUR FÜR
ANWENDUNGSORIENTIERTE FORSCHUNG UND INNOVATION
CHARLOTTE ALBER
- 21 THE FUNDING SCHEME “IMPACT INNOVATION” – THE IDEA
BEHIND AND THE RESULTS OF AN ASSESSMENT
PHILIPP AIGINGER-EVANGELISTI, BARBARA GLINSNER AND DOROTHEA STURN
- 27 FUNDING RESEARCH DATA INFRASTRUCTURES: FUNDING
CRITERIA IN GRANT PEER REVIEW
RICHARD HEIDLER
- 37 ANALYSIS OF THE INNOSUISSE START-UP TRAINING IN A
THEORY OF PLANNED BEHAVIOR FRAMEWORK
ELISABETH NINDL AND PETER KAUFMANN
- 44 USING THE DELPHI METHOD IN EVALUATIONS –
INCORPORATING A FUTURE ORIENTED PERSPECTIVE
IN EVALUATIONS
MICHAEL DINGES, ANNA WANG AND KLAUS SCHUCH

- 51 WIE KÖNNEN WIDERSPRÜCHE ZWISCHEN
MARKTDURCHDRINGUNG UND REBOUND-VERMEIDUNG
GELÖST WERDEN?
SYSTEMANALYSE KONVERGENTER UND DIVERGENTER EINFLUSSFAKTOREN AN
DEN BEISPIELEN E-AUTO UND GEBÄUDESANIERUNG
VERONIKA KULMER, SEBASTIAN SEEBAUER UND CLAUDIA FRUHMANN
- 60 FAIR PROJECTS – BAD DATA?
EVALUATING THE GENDER BALANCE IN SCIENCE PROJECTS
MAGDALENA KLEINBERGER-PIERER, SIMON POHN-WEIDINGER AND KARIN GRASENICK
- 72 EINE UNTERSUCHUNG DER MARKTSITUATION IM BEREICH
DER FTI-EVALUIERUNG IN ÖSTERREICH
JÜRGEN STREICHER, WOLFGANG POLT UND MAXIMILIAN UNGER

EDITORIAL

DEAR READERS!

TIME GOES BY.

With this issue we are celebrating the **50th anniversary of the *fteval Journal for Research and Technology Policy Evaluation***.

The first thirteen issues were published under the title *'Platform Technology Evaluation'*. Different founding members of the platform acted as editors of individual issues, such as **Oliver Fritz, Wolfgang Polt, Gernot Hutschenreiter, Michael Stampfer, or Anton Geyer**. The journal understood itself as an open, irregularly published forum for discussion of methodological and content-related evaluation issues in the field of research and technology policy. And this idea is still the basic direction of the journal. From the very beginning it invited and accepted contributions from scholars from all over the world.

As of number 14, the journal was newly labelled as *'Newsletter of the Austrian Platform for Research and Technology Policy Evaluation'*. **Klaus Zinöcker** became the driving force behind the journal until he handed over to myself. In 2013, finally, the Journal got a new refurbishment. It did not only receive a new design, but also its current name. Moreover, an international editorial review board was introduced that became responsible for the quality assurance of the submitted papers. It currently consists of **Rebecca Alinson, Balázs Borsi, Elke Dall, Michael Dinges, Leonid Gokhberg, Wolfgang Polt, Andreas Reinstaller, Klaus Schuch, Michael Stampfer** and **Lena Tsipouri**. The members of the editorial board ensure that contributions fit the subject area, do not contravene any of our policies, are of good quality, and that they can reasonably be considered a part of the practice-oriented academic literature. Still today, the *'fteval Journal for Research and Technology Policy Evaluation'* works at the interface between practice-oriented research and technology policy and academic quality. The journal calls for contributions from policy makers, policy-delivery systems, evaluators and the academic community. It also addresses these communities.

The journal is **open access** and it does not charge fees to authors for contribution or processing ('Diamond Standard'). Each journal issue and each paper published thereunder receives today a separate DOI (recorded under crossref). Authors or their institutions retain copyright to their publications. All articles are published under the Creative Commons Attribution license. The journal is also registered in OpenDOAR.

The journal at hand informs about recent evaluations and evaluation methods in the domain of R&I policies. I am thankful to the contributors of this issue and to all contributors and supporters of the 49 issues which we could publish before.

AND THE STORY GOES ON.

If you want to publish in the *'fteval Journal for Research and Technology Policy Evaluation'*, please drop me an email or consult our website to learn more about the journal, its regulations and opportunities for publishing!

Yours sincerely

Klaus Schuch

Executive Manager of the Austrian Platform for Research and Technology Policy Evaluation

March 2020

MONITORING NATIONAL INNOVATION PERFORMANCE: GOING BEYOND STANDARDIZED RANKINGS

JOHANNES GADNER AND JÜRGEN JANGER

DOI: 10.22163/fteval.2020.463

INTRODUCTION¹

Innovation performance has become increasingly important for governments as they search for ways to stimulate the economy and to address pressing societal challenges (e.g. Edler/Fagerberg 2017; Androsch/Gadner 2015; Aghion/David/Foray 2009). Thus, in recent years a variety of innovation performance rankings have been developed to measure performance levels and benchmark them against other countries. Those are, hence, closely watched by policymakers and are frequently perceived as a neutral gauge of a country's innovation performance (Schibany/Streicher 2008). However, standard rankings such as the *European Innovation Scoreboard* (EIS) or the *Global Innovation Index* (GII) in general apply a standardised set of indicators to all countries ranked without taking account of country specific weaknesses or strengths. For example, the Austrian government has been using the benchmark defined by the EIS to compare the Austrian innovation performance in the leading countries in Europe. Yet, the EIS should not be used as the only criterion to assess the effects of innovation efforts due to well-known shortcomings which result both from its construction as a composite indicator and from problems related to the substance of the indicators which make up the EIS summary index.

Concerning the latter, it underestimates e.g. the effects of innovation activities because it does not include indicators for improvement in the existing sectors, i.e. intra-sectoral structural change or sectoral upgrading; see also the more detailed explanations further below in this paper (Janger et al. 2017). Several of the other indicators used may lead to misleading interpretations. By way of example, the indicator measuring the share of knowledge-intensive services exports in total services exports penalizes countries with a large tourism sector such as France or Austria; some of the indicators drawn from the Community Innovation Survey are quite volatile or yield surprising results, e.g. that the share of new to market innovations (a proxy for novelty) is higher in some Eastern European countries than in Germany.

On top of these problems of substance, the construction of the EIS also entails drawbacks. First, it aggregates input and output indicators, so that ample inputs can compensate for weak outputs. (e.g. Edquist et al. 2018). Second, composite indicators are prone to the discussion

on appropriate weighting of the indicators – should all indicators carry the same weight in determining innovation performance or are some more important than others? Grupp and Schubert (2010) have pointed out that different weighting schemes can give rise to very different country rankings. They recommend complementing composite indicators by multidimensional representations of the issues leading to the performance in the aggregate indicator, i.e. explaining to policy makers why the performance arises. This would safeguard the communication function of composite indicators which due to their simplicity (a simple country rank) can easily be used for drawing the attention of a wider audience to problems of innovation performance.

In this article, we sketch such a multidimensional framework. It is applied by the Austrian Council for Research and Technology Development – the central advisory body of the Austrian Government for education, science, research and innovation policy affairs – for monitoring, measuring and evaluating the performance of Austria's national innovation system in international comparison. The paper aims to present an overview of the development of the measurement framework in co-operation with the Austrian Institute of Economic Research and in attunement with the responsible ministries. The article also intends to describe its application in the context of the Council's annual Reports on Austria's Scientific and Technological Capability.

The applied measurement framework for innovation performance takes into account country-specific conditions and provides information on both the current distance to strategic benchmarks as well as information on the distance to the benchmark at a given time horizon, based on extrapolating past growth trends. A matrix composed of the juxtaposition of current and future to target provides information able to guide policy priority setting, in terms of a measure of the effort required to reach targets (or the likelihood of reaching targets). Using this approach, the Austrian Council has been providing a sound source of evidence for an international comparison of Austria's innovation performance which supports evidence-based policy-making.

¹ A slightly modified version of this article will be published in Cele, M. / Luescher, T. / Wilson, F. A (Eds, forthcoming): Developing Innovation Excellence at a Global Level. HSRC Press, Pretoria.

TOWARDS A FRAMEWORK TO MEASURE INNOVATION PERFORMANCE

On 31 August 2009, the Austrian Council of Ministers agreed to set up an inter-departmental working group to elaborate a *Strategy for research, technology and innovation* (RTI Strategy). The aim was to define strategic goals and measures for Austrian RTI policy for the period up to 2020. The RTI Strategy with the title *Becoming an Innovation Leader* was adopted in March 2011. In its title, the document reflects the priority goal to be reached by 2020.² The government has commissioned the Austrian Council for Research and Technology Development with the task of measuring the progress of the implementation of the strategy and monitoring Austria's performance in RTI compared to that of the leading countries in Europe and the world. Since then, the Council has been drawing up a yearly Report on Austria's scientific and technological capability³ to present the main findings of its monitoring activities. The first report was presented in June 2012.

The reports apply a framework to measure innovation performance that relies upon a thorough analysis of the RTI system, strategic goals set by the government, and standardised indicators to operationalise monitoring of target achievement. In a first step, a detailed country-specific analysis of strengths and weaknesses of Austria's innovation performance was conducted, which was intended as groundwork for the subsequent definition of strategic goals compiled in a comprehensive RTI Strategy. This multi-year process of discussions and analyses consisted of three pillar stones:

- The "Austrian Research Dialogue"⁴ (2007-2008) was designed to be a broad, nationwide process of discourse and consultations with Austrian stakeholders for the purpose of further developing the innovation system and a knowledge-based society.
- The evaluation of Austria's research funding⁵ ("System Evaluation" 2008-2009) provided a profound assessment

of the entire competitive R&D funding system, along with recommendations for improvement by experts.

- The Council for Research and Technology Development elaborated evidence-based strategic proposals and recommendations for further development of the Austrian research and innovation system ("Strategy 2020" 2009).⁶

Based upon the results of these analytic processes, strategic goals for improving the Austrian innovation system were adopted in the aforementioned RTI Strategy and five performance areas (education, tertiary education, academic research, corporate R&D as well as governance and funding of R&D) were identified with particular catch-up potential. This is where our monitoring framework comes in: While some quantitative targets were set for each performance area (for example the goal to reach an R&D quota of 3.76 percent by 2020), most goals with respect to the performance areas were only expressed in a qualitative way, so that we constructed targets from the average values of the group of *Innovation Leaders* (according to the *European Innovation Scoreboard* (EIS) at which the RTI Strategy is oriented).⁷ This means that some of our indicators use as target value the number which was already set in the RTI Strategy, while most use as target value the arithmetic average of the group of leading innovation countries as determined by the EIS.

In fact, any target value can be used; while an absolute value will be static, the average of a number of freely chosen peer countries will be dynamic in the sense that the target value changes according to the performance of the peer countries. This is similar to standardised rankings, with the exception that for our tailor-made approach peer countries can be chosen freely, presumably from a set of countries whose structures and performances are not too far away from the country to be monitored.

To construct targets, appropriate indicators had to be selected, in terms of relevance and reliability, but also data availability. For the purpose of measuring Austria's innovation performance, a set of 75 indicators were developed in co-operation with the Austrian Institute of Economic Research and discussed with the responsible ministries.⁸ The selected indicators are based on internationally used classifications of

2 The priority goal of the Strategy to become an Innovation Leader by 2020 is oriented at the European Innovation Scoreboard (EIS). As a Strong Innovator Austria currently ranks among the top 10 member states of the European Union. Austria's score amounted to 120 points according to EIS 2018, while the group of Innovation Leaders reached an average score of 135. For details see https://ec.europa.eu/growth/industry/innovation/facts-figures/scoreboards_en

3 All reports can be downloaded from <https://www.rat-fte.at/performance-reports.html>

4 The Research Dialogue was initiated by the Ministry of Science and Research. A synopsis (in German) can be downloaded here: <https://bmbwf.gv.at/forschung/national/standortpolitik-fuer-wissenschaft-forschung/oesterreichischer-forschungsdialog/>

5 The "Evaluation of Government Funding in RTDI from a Systems Perspective in Austria" was commissioned by the Ministry of Transport, Technology and Innovation. The Synthesis Report (in English) can be downloaded here: https://www.bmvit.gv.at/service/publikationen/innovation/forschungspolitik/downloads/systemevaluierung/synthesis_report.pdf

6 The "Strategy 2020" was elaborated by the Austrian Council with the support of a Web-based discussion platform in order to deepen and intensify the process through virtual interaction with all relevant stakeholders and the RTI community. For the community-based innovation approach for strategy development see Gardner/Leo 2010. The Strategy 2020 (in English) can be downloaded here: <https://www.rat-fte.at/strategy-2020.html>

7 The term *Innovation Leader* refers to those EU countries in the top group in the annual *European Innovation Scoreboard* (EIS) of the European Commission. Today, the group comprises Denmark, Germany, Finland, Luxembourg, the Netherlands and Sweden. For the calculations the "innovation leaders' actual value" is used for every performance area.

8 It has to be mentioned that the processes of country analysis, target-setting and indicator choice enable collective learning and discussion processes at the national level, something which is completely absent from standardised innovation rankings but fosters the legitimacy of S&T indicators (Barré 2010). In our case six ministries were involved in the process: The Federal Chancellery (BKA), the Ministry of Finance (BMF), the Ministry of Education, Art and Culture (BMUKK), the Ministry of Transport, Innovation and Technology (BMVIT), the Ministry of Science and Research (BMWV) and the Ministry of Economic Affairs, Family and Youth (BMWFJ).

the OECD, Eurostat etc. and corresponding data portfolios. These are collected regularly on a national as well as on an international level.⁹ Every indicator corresponds to a strategic target of the RTI Strategy to operationalise goal achievement.¹⁰

For the development of the set of indicators it was important to focus on indicators suitable for better representing the conditions of the Austrian RTI system than for example those used by the EIS or other standardised rankings. This was deemed to be important by the Council since the EIS strongly focuses on inter-sectoral structural change and captures economic effects of innovation mainly by the growth of high-tech sectors; in contrast to this, the EIS underestimates effects of innovation activities in medium-tech-sectors, as intra-sectoral upgrading—improving innovation within a sector, rather than growing the share of innovation-intensive sectors in total output—is barely captured (Janger et al., 2017).

In fact, the EIS includes only few indicators for improvement across all existing sectors, i.e. intra-sectoral structural change or sectoral upgrading in less R&D intensive sectors such as steel or automobile parts. Austria performs very well in less R&D intensive sectors, which is why a picture not taking this aspect into account is rather incomplete. Moreover, the EIS indicators which could be interpreted as indicators of sectoral upgrading are based on the *Community Innovation Survey* (CIS).¹¹ As the CIS indicators are very volatile, their explanatory power is limited. Statistically more stable indicators such as the export quality in technology-oriented sectors or R&D intensity adjusted for the structural composition of an economy demonstrate only a relatively small gap to the leading innovation countries (for details see Austrian Council for Research and Technology Development 2014, pp. 10ff). Consequently, the indicators used within the framework adopted by the Austrian Council rely on more suitable indicators representing Austria's country-specific characteristics. This does not prevent international comparison as these indicators are also available for other countries and in fact, the whole framework rests on international comparison with leading countries.

VISUALISING INNOVATION PERFORMANCE

The applied framework focuses on the degree to which the goals set out in the strategy have been achieved (static component, current **distance to the goal**) and on the degree to which the goals may be achieved in the future (dynamic component, **probability of achieving the goal**). While the current distance to the target simply reflects Austria's current performance relative to the target value (either as set by the RTI strategy or as the average level of the *Innovation Leaders*), the probability of achieving the goal extrapolates past growth trends to indicate where Austria's performance would be by the time horizon 2020. This is intended to guide policy priority setting and assumes a "business as usual scenario", i.e. that if no more policy interventions happen, trends continue as in the past. Extrapolating past growth trends is not used to forecast indicator values in the future, but to show the development over time of an indicator as an additional information for policy purposes—see below the description of the four quadrants of figure 1.

This can be graphically displayed as the following example shows. Figure 1 provides a comprehensive overview of the performance within the analysed areas of the RTI system in relation to the selected peer countries which are as outlined above as those countries identified as "Innovation Leaders" by the EIS. This relates to the stated objective of the RTI strategy to catch up to the *Innovation Leaders* and is hence a policy choice rather than a choice based on an analysis of e.g. structural similarity.

The distance to target on the horizontal axis in figure 1 illustrates the current Austrian value and the distance to the respective target. It shows the ratio and the distance of the last available Austrian value to the national set target according to the RTI Strategy and the *European Education and Training 2020 Strategy*.¹² If there is no target value, the target is constructed out of the last available average value for the current Innovation Leaders according to the EIS. This is due to the fact that catching up with the group of *Innovation Leaders* is a priority goal for Austria, as established in the RTI Strategy. The arrows show the movement of indicator values over time, with the origin of the arrow denoting the indicator value in the first measurement year and the end of the arrow pointing to the current value.

9 A detailed explanation of the indicators (including the strategic goals which they operationalise) as well as the underlying data and the calculations can be found in the Appendices of every *Report on Austria's scientific and technological capability*.

10 It has to be noted, that quantitative indicators cannot display all the objectives of the RTI strategy. This particularly applies to the performance area of *governance structures of the RTI system*. Thus, the indicator-based analysis has always been complemented with a qualitative evaluation of implemented measures and goal achievement.

11 The *Community Innovation Survey* (CIS) is a survey of innovation activity in enterprises. It is carried out biennial by EU member states and number of ESS (European Statistical System) member countries. Although it is designed to provide information on the innovativeness of sectors by type of enterprises, on the different types of innovation and on various aspects of the development of an innovation, such as the objectives, the sources of information, the public funding, the innovation expenditures etc, the compiling of the data is voluntary to the countries. This means that in different surveys or years different countries are involved. For details see <https://ec.europa.eu/eurostat/web/microdata/community-innovation-survey>

12 Following the suggestion of the Federal Ministry of Education, Arts and Culture at the time, target values of the *European Education and Training 2020 Strategy* have been used for some indicators operationalising targets within the education system.

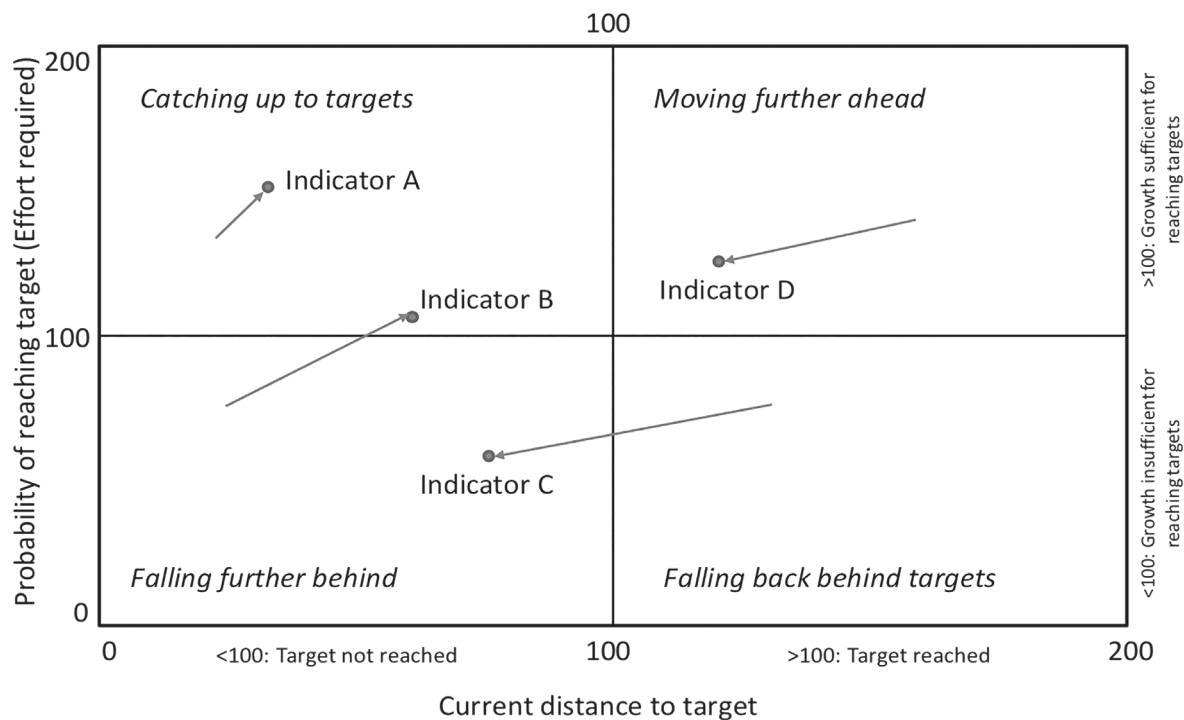


Figure 1: Visualising innovation performance: Outcomes over time based on the combinations of current and future distance to target
 Source: Austrian Council for Research and Technology Development (2018): Report on Austria's Scientific and Technological Capability 2018. Vienna, p. 103.

The distance to target exhibits Austria's current performance level and reveals the current difference between performance achieved and target performance, but it says nothing about the changes or dynamics which are required for the goal to be achieved. Thus, an indicator which currently lies just below the target level could therefore deteriorate again due to a negative development. To put it another way, the exclusive comparison of the distances to target does not allow any conclusions with regard to the prospect of goal achievement. For this reason, the probability of reaching the target on the vertical axis was selected as the second dimension. It shows whether or not the past growth of the indicator is sufficient for goal achievement. It indicates the ratio of the average annual growth rate of the respective data series in the past and the projected value for Austria in 2020 to the target value for 2020 (under a business-as-usual assumption, see above). As target value for the calculation of growth, if no quantitative target value was set in the RTI strategy for the respective goal, the projected value for 2020 of the *Innovation Leaders* will be used, and not their actual value. This in turn is determined on the basis of the average growth rates of the comparison countries in the past.

All indicators can be interpreted in the same direction, i.e. values above 100 signal goal achievement, values below 100 a corresponding distance to the goal. Indicators with a probability of reaching the goal below 100 – i.e. below the horizontal line at 100 – are based on past growth trends unlikely to reach the target by the end of the time horizon; indicators with a value above 100 – above the horizontal line – are likely to reach the target value. The standardisation of the values is achieved as follows: The Austrian value is divided by the respective target value and multiplied by 100. If performance improvements are accompanied

by a decline in the indicator values, such as with the unemployment rate, for example, the values would be inverted (i.e. target value in the numerator, Austrian value in the denominator), in order to retain the interpretation of "greater equal to 100 = goal achievement". Values above 200 are limited in the graphics to 200.

If an indicator is located in one of the two quadrants on the left side, this means that Austria has not yet achieved the set goal. For indicators in the bottom left corner this will very probably also remain unchanged. Hence, due to the weak growth rate in these performance areas, Austria will not achieve the goal by 2020. Consequently, without additional measures the *Innovation Leaders* are very likely to remain in front. Measures that are suitable to increase performance in this area should therefore be handled as a special priority. Indicators in the top left area are catching up, which could result in achieving the goal by 2020, as the Austrian development dynamic is greater than that of the comparison countries. In these performance areas no further measures would be required, always assuming that trends go on as in the past.

Indicators located in the two right quadrants show that the corresponding goals have already been achieved. Indicators in the top right corner signify that Austria has achieved the goal and in all likelihood will also remain in front until 2020 due to the high growth rates, provided the growth of the comparison countries remains within the expected range. Thus, there is currently no need for further action. Indicators in the quadrant on the right bottom signal that Austria's growth is insufficient to maintain its edge ahead of the *Innovation Leaders* in the long term. Accordingly, the development should be observed very closely here to also counteract in good time where required.

ILLUSTRATING THE USE OF THE FRAMEWORK WITH A PRACTICAL EXAMPLE

To illustrate how the framework is used in practice, an example from the *Report on Austria's scientific and technological capability 2015* is presented below.¹³ Figure 2 shows the indicator-based results of the approach for the area of research and innovation in the corporate sector for the current distance to the target and the probability of reaching the target. The tail of each indicator depicts the past development from 2010 until 2015.

The figure reveals at a glance the developments of both the current distance to the goal and the probability of reaching the goal by 2020 in the performance area of research and innovation in the corporate sector within the Austrian RTI system. It becomes clear that at the given time there was still considerable potential to optimise performance and increase efficiency in the corporate sector. While the six indicators in the right upper quadrant had indeed already reached their target or the level of the Innovation Leaders, the majority of indicators still continued to be within the bottom left quadrant indicating a performance below the

average level of the peers and expected to fall further behind. Judging by the current trends, and with no additional measures taken the indicator in the bottom right corner will deteriorate and fall below the target level by 2020. Those indicators in the top left quadrant display dynamic developments and will most likely continue catching up and reach the target level by the year 2020.

In sum, the presented performance area of research and innovation in the corporate sector at the given time had the third-smallest goal distance of the six areas analysed and the highest probability to achieve the goal. Overall, goal distance was deemed to be within reach by 2020 – provided the trends continued to develop as they have in the past. However, many indicators, such as 'Competition policy' or 'Venture capital intensity', were well below the target level and displayed insufficient or retrograde dynamics to catch-up with the group of leading countries. Furthermore, performance in key sub-areas, such as start-up dynamics and venture capital intensity, remained extremely weak in comparison to the Innovation Leaders. Based on this picture, the relevant fields of action were addressed in the Council's report – including concrete policy recommendations for improving Austria's innovation performance within the specific areas.

It is important to mention that the Council of Ministers commissioned the Council for Research and Technology Development to elaborate the annual *Reports on Austria's scientific and technological capability* in

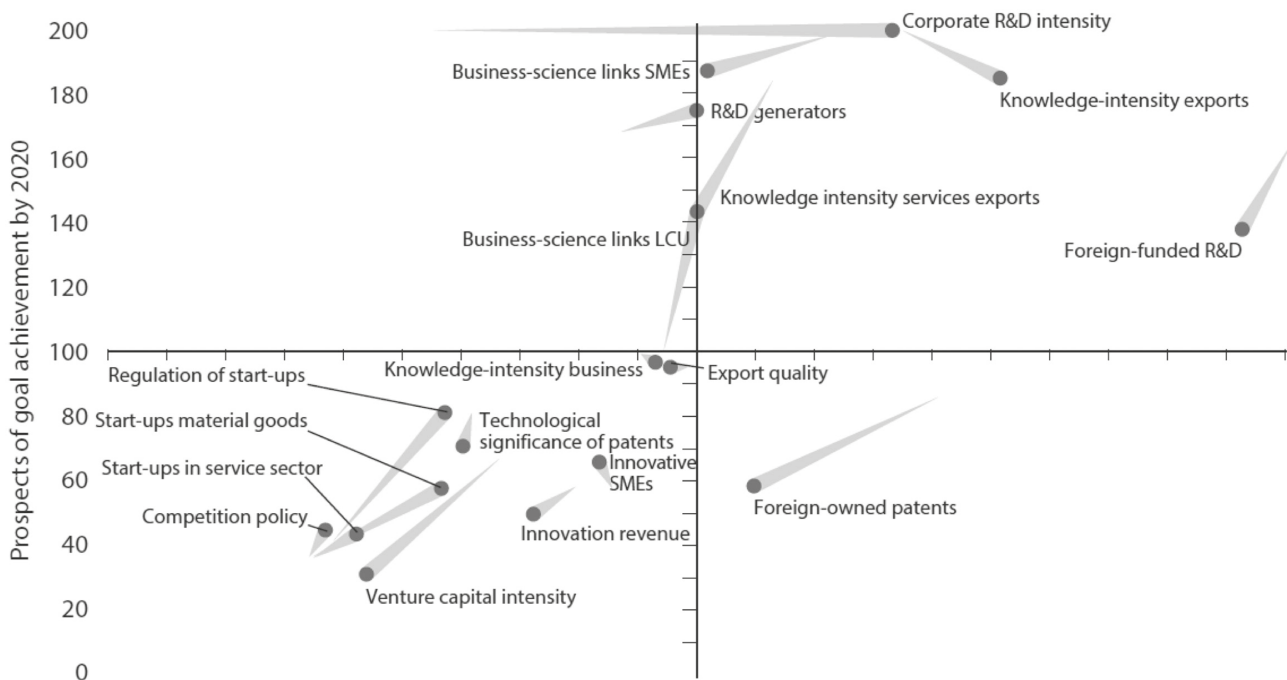


Figure 2: Development of distance to targets and probability to reach the target in the area of research and innovation in the corporate sector, 2010 to 2015
 Source: Austrian Council for Research and Technology Development (2015): *Report on Austria's Scientific and Technological Capability 2015*. Vienna, p. 47.

order to send it to the Austrian Parliament. There it is debated in the parliamentary Committee on Research, Technology and Innovation. It is also presented publicly and published on the Councils Web-page. With this approach it was intended to put the political and public debate on education, science, research and innovation policy affairs on a sound evidence-based footing.

CONCLUDING REMARKS

The described measurement framework is a sound basis for a meaningful monitoring of national RTI policies and innovation performance, although it requires more work than using standard indicators as outlined above: first, goals have to be set, possibly within a strategy-building exercise. Goal-setting requires an analysis of the strengths and weaknesses of a country. Once goals are set, suitable indicators can in a second step be drawn from available sources, such as OECD data; when no ready-made indicators exist, new indicators would have to be built, which needs more effort and above all the corresponding data. Once goals and indicators are in place, progress over time can be monitored, assessing the impact of any policies implemented to reach the goals.

As we have demonstrated, the sketched framework for measuring innovation performance not only refers to Austria's country-specific conditions but also facilitates a dynamic view of past and future developments. While standard rankings just give a static snapshot of one performance point in time, the presented approach yields the possibility to show where current dynamics of innovation performance may lead to in the near future. Growth trends of single indicators may be positive or negative so that a yearly snap shot on its own is of little use to policy-makers. Different policy priorities apply to areas, where a country lags behind, but catches up quickly as compared to those, where a country lags behind with a further falling tendency. In the first case, no further action is needed, in the second, alarm bells should be ringing.

The limitations of the proposed framework for policy makers are that the framework needs to be custom-made for national purposes and it requires more resources first for the analysis of national performance and then for national target-setting or strategy-building. Like a tailor-made suit, the framework fits better but is more expensive and time-consuming. The benefits of appropriately reflecting country strengths and weaknesses may however largely outweigh the costs of not only relying on standardised indicator frameworks. Moreover, common to all indicator-based measurement frameworks, it is a quantitative framework, meaning that international benchmarking relies on available data for indicator-building. Country-specific challenges lacking internationally comparable data can only be addressed in a qualitative way.

For the Austrian Council, the described monitoring framework and the indicator-based measurement of innovation performance has been a basis from which to draw conclusions about Austria's scientific and technological capability in relation to the leading innovation nations. With this evidence-based analysis of strengths and weaknesses of the Austrian RTI system, the most pressing priority fields of action have been identified and dealt with in the Council's annual reports and policy recommendations. Additionally, the Council uses the described approach for the strategic monitoring of the realisation of the federal government's RTI strategy as well as the assessment of efficiency and effectiveness

of the implemented measures. The measurement framework can also be adopted in other countries, provided that there are suitable strategy-building and performance-analysing exercises. It may be an interesting tool for policy-makers elsewhere, allowing them to focus on country-specific issues rather than on insights based on standardised innovation rankings.

REFERENCES

- Aghion, P., David, P. A. and Foray, D.** (2009): Science, technology and innovation for economic growth: Linking policy research and practice in "STIG Systems", *Res. Policy*, 2009, 38(4), pp. 681–693.
- Androsch, H. and Gadner, J.** (2015): Designing the future. In: Austrian Council for Research and Technology Development (ed.): *Designing the future: Economic, societal and political dimensions of innovation*. Echo-media, Vienna, pp. 8-39.
- Austrian Council for Research and Technology Development** (2018): Report on Austria's Scientific and Technological Capability 2018. Vienna.
- Austrian Council for Research and Technology Development** (2015): Report on Austria's Scientific and Technological Capability 2015. Vienna.
- Austrian Council for Research and Technology Development** (2014): Report on Austria's Scientific and Technological Capability 2014. Vienna.
- Barré, R.** (2010): Towards socially robust S&T indicators: indicators as debatable devices, enabling collective learning. *Res. Eval.*, 2010, 19(3), pp. 227–231
- Edler, J. and Fagerberg, J.** (2017): Innovation policy: what, why, and how. *Oxf. Rev. Econ. Policy*, 2017, 33(1), pp. 2–23.
- Edquist, C., Zabala-Iturriagoitia, J. M., Barbero, J. and Zofio, J. L.** (2018): On the meaning of innovation performance: Is the synthetic indicator of the Innovation Union Scoreboard flawed? *Research Evaluation*, 2018, 27(3), pp. 196–211
- Gadner, J. and Leo, H.** (2010): Innovation Policy Counseling 2.0: How Open and Community-Based Innovation Processes Promote Political Decision Making. *bridges*, vol. 28, December 2010 / Feature Article. Online: <https://ostaustria.org/bridges-magazine/volume-28-december-20-2010?start=20>
- Grupp, H. and Schubert, T.** (2010): Review and new evidence on composite innovation indicators for evaluating national performance. *Res. Policy*, 2010, 39(1), pp. 67–78.
- Janger, J., Schubert, T., Andries, P., Rammer, C. and Hoskens, M.** (2017): The EU 2020 innovation indicator: A step forward in measuring innovation outputs and outcomes? *Res. Policy*, 2017, 46(1), pp. 30–42.
- Schibany, A. and Streicher, G.** (2008): The European Innovation Scoreboard: drowning by numbers? *Sci. Public Policy*, 2008, 35(10), pp. 717–732.

AUTHORS

JOHANNES GADNER

Austrian Council for Research and Technology Development

Pestalozzigasse 4/D1, 1010 Vienna

<https://www.rat-fte.at/home-en.html>

E: j.gadner@rat-fte.at

JÜRGEN JANGER

Austrian Institute of Economic Research

Arsenal, Object 20, 1030 Vienna

<https://www.wifo.ac.at/en>

E: juergen.janger@wifo.ac.at

KEYWORDS:

monitoring, innovation performance, RTI policies, measurement framework, international comparison

ZUGANG ZU REGISTER- UND INDIVIDUALDATEN FÜR DIE WISSENSCHAFTLICHE FORSCHUNG IN ÖSTERREICH

GERHARD SCHWARZ, THOMAS KÖNIG, HARALD OBERHOFER UND MICHAEL STRASSNIG

DOI: 10.22163/fteval.2020.464

1. EINLEITUNG

Eine zentrale Aufgabe öffentlicher Verwaltung besteht darin, Datenbestände anzulegen – zur Erfüllung ihrer ureigensten Aufgaben, aber auch zur Nutzung durch andere. Diese Datenbestände dienen als Entscheidungsgrundlage, zur Planung, Dokumentation, Kontrolle und Evaluierung. So entstehen Register: von der Todesursachenstatistik über die Sozialversicherungsdaten und die Transparenzdatenbank bis zur Statistik Austria, dem größten österreichischen „Datenbunker“.

Der Zugang zu vielen hochwertigen Datensätzen ist in der Regel stark eingeschränkt, oder gar nicht möglich.¹ Dabei geht es um die Wahrung des Datenschutzes bzw. der Geheimhaltungsbedürfnisse der in den Registern erfassten natürlichen und juristischen Personen. Es liegt aber auch nicht unbedingt im Eigeninteresse der Bürokratie, Daten nach außen zu geben, die gegebenenfalls politisch sensibel sind. Ein besonders anschauliches Beispiel publizierte die New York Times kürzlich in Bezug auf ihre Recherchen zu den EU-Agrarsubventionen:

The European Union, we learned, maintains a database with all this information. But when we requested it, officials responded both that the data did not exist and that it did exist but could not easily be extracted. When we pointed out that the European Union had previously managed to extract the information for a World Bank study, they said the data was private. We appealed. We were denied again. Downloading the information from the government's own computers was so complicated, officials said, that they could not be forced to do it. (Apuzo 2019)

Jedoch: Hat die Öffentlichkeit nicht das Recht, über die Verwendung von staatlichen Mitteln Bescheid zu wissen? Und ist es nicht die Pflicht

einer öffentlichen Verwaltung in einer demokratischen Gesellschaft, ihre Datensätze zur Analyse auf Basis von wissenschaftlichen, nachvollziehbaren Kriterien bereitzustellen, um evidenzbasierte Aussagen über staatliches bzw. hoheitliches Handeln zu erlauben und damit zu verbessern? Dies wäre zumindest im Rahmen einer politischen Diskussion über Nutzen und Risiken der wissenschaftlichen Verwendung von solchen Daten zu berücksichtigen – eine Diskussion, die aber bisher noch nicht geführt wurde. In einigen europäischen Ländern ist dies inzwischen auch passiert, und wenn auch die Grenze von Land zu Land verschieden gezogen wurde, so gilt für diese Fälle, dass sie dazu eine gültige Rechtslage hergestellt haben. Auch in Österreich hat sich dazu in den letzten Jahren etwas getan. Der vorliegende Artikel liefert dazu einen konzisen Überblick.

2. DATENZUGANG – STAND DER DINGE IN ÖSTERREICH

Zunächst: Die Hürden, mit denen die Journalisten der New York Times in Brüssel zu kämpfen hatten, sind den Wissenschaftlerinnen und Wissenschaftler in Österreich nur zu vertraut. Ein Beispiel betrifft die derzeit in Diskussion befindliche Forschungsfinanzierungsdatenbank (FFDB). Hier gibt es Vorbehalte, der Wissenschaft Zugang zu pseudonymisierten Daten der staatlichen Forschungsförderung auf Unternehmensebene zu gewähren.

Wie politisch heikel der Zugang der Wissenschaft zu öffentlichen Datenbeständen gesehen wird, zeigt auch das Forschungsorganisationsgesetz (FOG). Mit dem „Datenschutz-Anpassungsgesetz 2018 – Wissenschaft und Forschung“² wurde im FOG ein Rechtsanspruch wissenschaftlicher Einrichtungen auf Bereitstellung von Daten aus bundesgesetzlich vorgesehenen Registern geschaffen³. Dies war ein großer Erfolg des Bundesministeriums für Bildung, Wissenschaft und Forschung (BMBWF),

1 Eine Ausnahme bilden Informationen des öffentlichen Sektors (PSI), die bereits über Open Data Österreich (<https://www.data.gv.at/>) frei zugänglich sind. Oft handelt es sich dabei um Datensätze, die eine bestimmte Anwendungsrelevanz haben, aber für eine wissenschaftliche Nutzung nur von beschränktem Wert sind.

2 BGBl. I Nr. 31/2018

3 Einen Überblick zu den Vorzügen und Herausforderungen der Nutzung administrativer Daten in den Sozialwissenschaften bieten Connelly-Dibbend-Gayle-Playford. (2016).

das sich vehement für die Bereitstellung von Registerdaten für die Wissenschaft eingesetzt hat. Wie umstritten dieses Gesetz war, zeigen die knapp 80 Stellungnahmen, die für die parlamentarische Begutachtung von vielen Stakeholdern eingegangen sind. Neben Kritik am Gesetzesentwurf wurde insbesondere von der Wissenschaftsseite auf die Wichtigkeit für den Forschungsstandort hingewiesen.⁴

Leider kommt dieser wichtige Erfolg für die Wissenschaft mit zwei Wermutstropfen. Erstens haben spezifischere Regelungen zum Datenschutz und zur wissenschaftlichen Datenverarbeitung in Materiegesetzen Vorrang vor den Bestimmungen des FOG. Praktisch führt dies zu komplexen rechtsdogmatischen Diskussionen über die anzuwendenden Regelungen. So deutet die derzeitige Rechtsprechung der Datenschutzbehörde (DSB) darauf hin, dass sie – anders als die Erläuterungen zur Regierungsvorlage zu den einschlägigen Bestimmungen des FOG – keinen Vorrang der Bestimmungen des FOG gegenüber jenen des Datenschutzgesetzes (DSG) sieht. Endgültig entscheiden werden dies erst die Höchstgerichte.

Zweitens müssen Register, wenn die Bestimmungen des FOG anwendbar sind, erst einzeln mittels einer gemeinsamen Verordnung der/des zuständigen Ressortministers/in und des/der Bundesministers/in für Bildung, Wissenschaft und Forschung bereitgestellt werden. Damit bleibt es weiterhin den Ressorts vorbehalten, zu entscheiden, ob und welche bundesgesetzlich vorgesehenen Register der Wissenschaft zur Verfügung stehen. Auch angesichts der Vielzahl von für die Forschung relevanten Registern wird dies ein langwieriger Prozess sein.

3. VERHEDDERT IM REGELUNGSDSCHUNDEL

So wichtig die Implementierung der Registerforschung im FOG für die Wissenschaften in Österreich und ihren Beitrag zur „good governance“ auch ist, so ist es damit leider nicht gelungen, eine einheitliche Regelung des Zugangs der Wissenschaft zu den Mikrodatenbeständen der öffentlichen Hand zu schaffen. Register, die unmittelbar auf EU-Verordnungen oder Landesgesetzen beruhen, sind nicht vom FOG erfasst, das sich schon aus Kompetenzgründen auf bundesgesetzlich vorgesehene Register beschränkt. Ob und wann auch die Bundesländer ihre Register für die Wissenschaft öffnen werden, steht in den Sternen.

Parallel zur Regelung der Registerforschung im FOG bestehen bzw. entstehen weitere Regelungskorpusse für große und forschungsrelevante österreichische Datenbestände. So wurde im Transparenzdatenbankgesetz (TDBG) die Möglichkeit geschaffen⁵, Daten aus der Transparenzdatenbank, die die Einkommen und erhaltenen öffentlichen Leistungen aller natürlichen und juristischen Personen in Österreich enthält, „in anonymisierter Form an fachlich geeignete Personen oder wissenschaft-

liche Einrichtungen für wissenschaftliche Zwecke nach Vereinbarung der konkreten Anwendungsbereiche zu übermitteln“ (§ 34 Abs. 2 TDBG). Bemerkenswert daran ist einerseits, dass der Bundesminister für Finanzen nach dem Wortlaut der Bestimmung zwar verpflichtet ist⁶, die Daten für wissenschaftliche Verwendung zu übermitteln, allerdings nur zur Erfüllung der in § 2 TDBG genannten Zwecke⁷. Ob ein solcher Anspruch in der Praxis tatsächlich rechtlich durchsetzbar wäre, bleibt fraglich.

Andererseits ist an der Regelung der Datenübermittlung für wissenschaftliche Zwecke im TDBG bemerkenswert, dass sich diese nur auf „anonymisierte“ Daten bezieht. Somit ist der Zugriff der Wissenschaft auf grob auflösende Daten beschränkt, da ansonsten die Identifikation einzelner Merkmalsträger anhand bestimmter Ausprägungen von Merkmalen in den Daten nicht ausgeschlossen werden kann⁸. Es bedeutet jedoch auch, dass die Daten keine Schlüsselvariablen enthalten dürfen, die eine Verknüpfung mit anderen Datenbeständen ermöglichen würde⁹, da solche Daten nur „pseudonymisiert“, aber nicht „anonymisiert“ sind. Die Verknüpfung von Daten aus der Transparenzdatenbank mit anderen Datenbeständen zu Zwecken der Auswertung bleibt ausschließlich der „Statistik Austria“ vorbehalten (§ 18 Abs. 3 TDBG). Grob auflösende Daten und die Nichtverknüpfbarkeit sind aber ein Problem für hochqualitative und innovative Forschung, da mit solchen Daten heutzutage kaum mehr publizierbare Artikel zu produzieren sind.

Neben der Transparenzdatenbank ist gerade der Regelungsrahmen der bereits erwähnten Forschungsfinanzierungsdatenbank im Diskussion. Hier sollen alle Forschungsfinanzierungen durch die öffentliche Hand abrufbar sein, auch jene, die an private Unternehmen ausgeschüttet werden, einschließlich der steuerlichen Forschungsprämie. Nach gegenwärtigem Stand sollen der Wissenschaft nur zusammengefasste Daten bereitgestellt werden, aber keine Einzelfalldaten, bei denen die Identifikation einzelner Förderungsempfänger nicht ganz ausgeschlossen werden könnte. Dagegen sollen von der öffentlichen Hand bestellte EvaluatorInnen von Forschungsförderungsmaßnahmen auf Basis einer konkreten Beauftragung im Einzelfall Zugang zu Einzelfalldaten erhalten, sofern es für die konkrete Fragestellung notwendig und sinnvoll erscheint. Die Entscheidung bleibt also bei der beauftragenden öffentlichen Einrichtung.

Vergleicht man Regelungen zu Registerdaten nach FOG, zu Daten aus dem TDBG und zu Daten aus der FFDB, so sieht man, dass der Gesetzgeber keine einheitliche Vorstellung davon hat, unter welchen Bedingungen die Wissenschaft Zugang zu Einzelfalldaten erhalten soll. Dies zeigt sich auch bei den Bemühungen, der Wissenschaft einen besseren Zugang zu den Daten der Statistik Austria einzuräumen.

4 https://www.parlament.gv.at/PAKT/VHG/XXVI/ME/ME_00010/index.shtml#tab-Stellungnahmen

5 BGBl. I Nr. 70/2019

6 „Die Möglichkeit des BMF die Daten aus der Transparenzdatenbank in anonymisierter Form an Dritte zu übermitteln, soll dahingehend präzisiert werden, dass dies verpflichtend vorgesehen wird, allerdings mit der Einschränkung für wissenschaftliche Zwecke. Es soll gleicher Zugang für alle wissenschaftlichen Einrichtungen bei voller Transparenz bezüglich Verwendung der Daten und der Ergebnisse gelten. [...]“ (AA-109 XXVI. Gesetzgebungsperiode, Begründung)

7 Informationszweck, Nachweiszweck, Steuerungszweck, Überprüfungszweck und Wirtschaftlichkeitszweck.

8 Vgl. Erwägungsgrund 26 der Verordnung (EU) 2016/679 des Europäischen Parlaments und des Rates vom 27. April 2016 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten, zum freien Datenverkehr und zur Aufhebung der Richtlinie 95/46/EG (Datenschutz-Grundverordnung)

9 Etwa bereichsspezifische Personenkennzeichen (bPKs), wie sie das E-Government-Gesetz vorsieht.

4. DIE DATEN DER STATISTIK AUSTRIA

Der größte Datenhalter der öffentlichen Hand in Österreich ist die Statistik Austria.¹⁰ Dort liegen Kopien eines großen Teils der Daten die auch in den Registern der öffentlichen Hand stehen - und noch viel mehr: Von der Konjunkturerhebung über die Leistungs- und Strukturstatistik der österreichischen Wirtschaft, der F&E Erhebung, den Erhebungen zum Einsatz von Informations- und Kommunikationstechnologien in Haushalten und Unternehmen, dem Gebäude und Wohnungsregister bis hin zum Krebsregister und der Todesursachenstatistik, um nur einige zu nennen.

Der verbesserte Zugang der Wissenschaft zu den Einzelfalldaten der Statistik Austria würde einen riesigen Datenbestand für die wissenschaftliche Analyse zugänglich machen. Bereits vor zehn Jahren gab es daher – ausgehend von VertreterInnen der Wissenschaft (vgl. Gaulhofer, 2009) – Bemühungen, den Zugang der Wissenschaft zu den Datenbeständen der Statistik Austria zu verbessern. Schon damals wollte man eine anstehende Novelle des Bundesstatistikgesetzes nutzen, die Einzelfalldatenbestände der Statistik Austria für die Wissenschaft zu öffnen. Seither haben sich zahlreiche Menschen aus der Wissenschaft, der öffentlichen Verwaltung und auch aus der Statistik Austria für dieses Ziel verwendet (zuletzt etwa Kocher und Badelt, 2019); vor drei Jahren mündeten die Bestrebungen in der Schaffung der „Plattform Registerforschung“.¹¹

2017 und 2018 schien die (nach wie vor erforderliche!) Novellierung des Bundesstatistikgesetzes jeweils knapp bevorzustehen, um dann doch wieder verschoben zu werden. Der damals etablierte Dialog zwischen Politik, Statistik Austria, Stakeholdern und Wissenschaft ist inzwischen wieder zum Erliegen gekommen und das Bundeskanzleramt als zständiges Ressort arbeitet exklusiv mit Statistik Austria an einer Lösung, ohne die betroffene Wissenschaft oder andere Stakeholder weiter einzubinden. Dem Vernehmen nach ist in Bälde mit einem neuen Vorschlag des Bundeskanzleramts zu rechnen, auch weil eine Novellierung des Bundesstatistikgesetzes im Regierungsübereinkommen vorgesehen ist. Die Knackpunkte sind der Rechtsanspruch auf Zugang der Wissenschaft zu den Einzelfalldaten der Statistik Austria und dessen praktische Durchsetzung, die Verknüpfbarkeit mit anderen Datenbeständen, insbesondere Registerdaten nach FOG sowie die Gewährleistung des Datenschutzes sowie die Kosten des Zugangs zu Einzelfalldaten der Statistik Austria.

5. AUSTRIAN MICRO DATA CENTER

Trotz der erzielten Fortschritte, namentlich durch die Ermöglichung der Registerforschung im FOG, hinkt Österreich in puncto Einzelfalldatenzugang der Wissenschaft immer noch weit hinter internationalen Vorbildern wie Dänemark, Schweden oder den Niederlanden her (vgl. Falk-Hözl-Oberhofer, 2015). Die österreichische Bundesregierung kommt

in ihren Forschungs- und Technologieberichten der Jahre 2018 und 2019 zum selben Schluss (siehe Polt, Streicher et al., 2018 und Ecker et al., 2019). Damit ist evidenzbasierte Politik zum Vorteil der Allgemeinheit mangels datengetriebener Analysen weiterhin nicht – oder nur sehr eingeschränkt – möglich.

Ein Plan zur Abhilfe ist unter dem Schlagwort eines sogenannten „Austrian Micro Data Center“ immer wieder in Diskussion. Es wurde vom ehemaligen fachstatistischen Generaldirektor der Statistik Austria nach Vorbild des dänischen Mikrodatenzentrums eingebracht. Mit dieser Einrichtung liefert die dänische Statistikbehörde jährlich Daten für rund 400 neue Forschungsprojekte, und setzt dabei umgerechnet etwa vier Millionen Euro um.¹² Da Dänemark bevölkerungsmäßig etwa halb so groß ist wie Österreich, wird deutlich, welches Potential für den heimischen Wissenschaftsstandort hier schlummert.

Das „Austrian Micro Data Center“ würde wie ein One-Stop-Shop funktionieren, der Wissenschaftlerinnen und Wissenschaftern Zugang zu den Einzelfalldaten der Statistik Austria ebenso ermöglicht wie zu den Registerdaten nach FOG. Zentral ist dabei die Verknüpfbarkeit von Statistik- und Registerdaten miteinander auf Ebene der erfassten Einheiten, also in der Regel auf der Ebene einzelner natürlicher oder juristischer Personen bzw. Unternehmen. Auch die Verschneidung mit externen Daten, etwa Befragungsdaten,¹³ soll möglich sein. Daraus ergäben sich neue Perspektiven, etwa auch für Sozialerhebungen und andere Umfragen. Solche Verknüpfungen verschiedener Datenbestände erfolgen mittels nicht rückführbarer Kennzahlen, um den Datenschutz zu optimieren. Der Zugang zu den Daten würde über einen gesicherten Fernzugriff möglich sein, sodass die Daten nicht heruntergeladen werden können und das „Austrian Micro Data Center“ niemals verlassen.

6. DATENSCHUTZ, GEHEIMHALTUNG, SENSIBILITÄT UND GESELLSCHAFTLICHE VERANTWORTUNG

Besonderes Augenmerk liegt beim Zugang der Wissenschaft zu Register- und Statistikdaten auf dem Datenschutz, insbesondere der Geheimhaltung von personen- oder unternehmensbezogenen Daten. In Bezug auf gesundheitsbezogene Daten reagieren Datenschützer und Oppositionsparteien besonders sensibel, wie in Österreich in Zusammenhang mit der Novellierung des FOG festzustellen war (vgl. etwa ede/apa, 2018) und jüngst in Deutschland in Bezug auf eine erweiterte Nutzung von Gesundheitsdaten für Forschungszwecke (dpa, 2019). Die betroffenen Personen scheinen dies weniger dramatisch zu sehen: In Folge der breiten öffentlichen Diskussion um den Zugang der Wissenschaft zu den Daten der Elektronischen Gesundheitsakte (ELGA) meldeten sich geschätzt etwa 5.000 Personen zusätzlich ab (WZ online, APA, 2018). Dies entspricht nicht einmal 0,06 Prozent der im Jahresdurchschnitt 2018 an-

10 Eigentlich Bundesanstalt „Statistik Österreich“ (vgl. § 22 Abs. 1 Bundesstatistikgesetz)

11 Siehe <https://www.registerforschung.at>

12 Davon ist circa ein Drittel staatliche Basisförderung, der Rest wird den Betreibern der durchgeführten Projekte in Rechnung gestellt.

13 Vgl. etwa die Möglichkeit in den Niederlanden Daten des „Survey of Health, Ageing and Retirement in Europe“ (SHARE) mit administrativen Daten zu verknüpfen: <http://www.share-project.org/special-data-sets/record-linkage-project/linkage-share-nl.html>

spruchsberechtigten Personen in der österreichischen Krankenversicherung, und ist auch nur ein Bruchteil jener, die sich in den Jahren davor bereits – aus anderen Gründen – von ELGA abgemeldet haben.¹⁴

Registerdaten und Einzelfalldaten der amtlichen Statistik sind in der Regel nicht vollständig anonymisierbar und das kann datenschutzrechtlich tatsächlich ein ernstzunehmendes Bedenken sein. Allerdings: Die im FOG vorgeschriebene Pseudonymisierung, also das Ersetzen von Namen durch nicht rückführbare Kennzahlen, die (soweit man weiß) auch beim zukünftigen Zugang der Wissenschaft zu Einzelfalldaten der Statistik Austria eingesetzt werden soll und die das Verknüpfen unterschiedlicher Datenbestände erlaubt, ist darauf ausgelegt, Personen und Unternehmen zu schützen, deren Daten in einem Register oder in der amtlichen Statistik erfasst sind. Dies gilt bis auf wenige Ausnahmen.

Um welche Ausnahmen kann es sich hier handeln? Dazu zwei (theoretische) Beispiele: So könnten Wissenschaftlerinnen und Wissenschaftler einzelne Personen, etwa nahe Verwandte oder gute Freunde identifizieren, von denen sie einige kritische Merkmalsausprägungen kennen. Auch Unternehmensdaten könnten dann betroffen sein, wenn es in einzelnen Branchen nur wenige große Unternehmen gibt, die ohne Anführen der Firma anhand der Kombination aus Branche und Beschäftigtenzahl identifizierbar wären.

Für diese (seltenen, aber nicht auszuschließenden) Ausnahmen liegt es am Gesetzgeber eine Entscheidung zu treffen. Wir sind der Meinung, dass der potentielle Nutzen der Gesellschaft an den Ergebnissen der Registerforschung höher zu bewerten ist als der mögliche Nachteil, den einzelne Personen erleiden könnten, würden sie über Registerdatensätze identifiziert. Dies aus drei Gründen. Erstens, weil das FOG ohnehin weitreichende Datensicherheitsmaßnahmen vor, um jeglichen Missbrauch zu verhindern (§ 2d Abs. 1 FOG). Außerdem fallen Wissenschaftlerinnen und Wissenschaftler, die den Datenzugang missbrauchen, unter den Strafraum der Datenschutz-Grundverordnung. Und auch alle bisher bekanntgewordenen Entwürfe zum Zugang der Wissenschaft zu Einzelfalldaten der Statistik Austria sehen rigide Sicherheitsmaßnahmen vor sowie massive Konsequenzen für Fehlverhalten bis hin zur gerichtlichen Strafbarkeit.

Zu den Sicherheitsmaßnahmen gehört es – zweitens – auch, dass im Fall des Austrian Micro Data Centers die Daten nur im Fernzugriff zugänglich sein sollen und dessen Server niemals verlassen würden. Dies würde auch für die Registerdaten nach FOG gelten, wenn sie über das Austrian Micro Data Center verfügbar gemacht werden. Diese strengen Vorgaben zu Sicherheitsmaßnahmen, die Pseudonymisierung und die hohen Strafdrohungen scheinen unseres Erachtens geeignet, das Risiko für Fehlverhalten und Datenmissbrauch effektiv auszuschalten. Dies ist zudem aus den Erfahrungen jener Länder abzuleiten, die der Wissenschaft bereits einen besseren Zugang zu ihren Register- und Einzelfalldaten ermöglicht haben.

Drittens ist hier darauf hinzuweisen, dass sich innerhalb der Wissenschaft die forschungsintrinsischen Maßstäbe verändert haben. Die Forschenden sind bezüglich ethischer Fragestellungen ebenso sensibler geworden, wie grundsätzliche Aspekte der Wissenschaftsintegrität thematisiert werden – und das impliziert die gesellschaftliche Verantwortung. So haben sich auch innerhalb der Wissenschaft Institutionen, Policies und Verfahren zur Klärung ethischer Problemstellungen etabliert,

die für die Forschende einen handlungsnormativen Rahmen festlegen. Dazu zählen etwa die Österreichische Agentur für Wissenschaftliche Integrität (OeAWi) oder die an zahlreichen Forschungseinrichtungen mittlerweile etablierten Ethikkommissionen.

7. KONKRETISIERUNG DES REGIERUNGSPROGRAMMS

Die neue Bundesregierung sollte aus Sicht der Plattform Registerforschung in der Konkretisierung der im Regierungsprogramm formulierten Vorhaben jedenfalls zu folgenden drei Punkten aktiv werden:

Erstens, der mangelhafte Zugang der Wissenschaft zu Einzelfalldaten der Statistik Austria ist ein hausgemachtes Problem. Dänemark, Schweden oder die Niederlande, in denen der Zugang zu Daten der amtlichen Statistik liberaler geregelt ist, unterliegen demselben europäischen Rechtsrahmen und können daher als Vorbild dienen. Der österreichische Gesetzgeber sollte den Forschenden in diesem Land einen Datenzugang auf internationalem Niveau ermöglichen und damit den Wissenschaftsstandort Österreich stärken.

Erfreulicherweise sieht das aktuelle Regierungsprogramm in diesem Bereich die Einrichtung eines Austrian Micro Data Center vor. Zweitens sollte dieses Center so gestaltet werden, dass es den einfachen und sicheren Zugang der Wissenschaft zu den Einzelfalldaten der Statistik Austria und der Registerdaten nach FOG in einem One-Stop-Shop ermöglicht. Im Zuge der Einrichtung des Centers sollten auch die Bundesministerinnen und Bundesminister ermutigt und aufgefordert werden, jene Registerdaten für die Wissenschaft zu öffnen, die in ihren Ressorts schlummern.

Die Bemühungen sollten sich nicht nur in der Schaffung des erforderlichen legislativen Rahmens erschöpfen, sondern den tatsächlichen Fortschritt laufend überprüfen. Oft sind gesetzliche Regelungen im ersten Wurf noch unausgegoren und müssen nachjustiert werden. Aber auch eine perfekte rechtliche Grundlage ist erst die halbe Miete; sie muss erst von den betroffenen Institutionen gelebt und umgesetzt werden. Dies betrifft die Statistik Austria ebenso wie die Bundesministerien und anderen Einrichtungen, die bundesgesetzlich vorgesehene Register führen. Das Regierungsprogramm bleibt hier auf der Ebene allgemeiner Bekenntnisse. Drittens ist daher als konkreter Schritt auch ein solches Monitoring der Umsetzung der Datenzugangsstrategie zu etablieren.

Mit diesen drei Maßnahmen wäre ein wesentlicher Schritt getan, den Wissenschaftsstandort Österreich abzusichern und zugleich auch wieder näher an jene Länder heranzuführen, die hier inzwischen eine internationale Vorbildfunktion einnehmen. Davon würden natürlich Österreichs wissenschaftliche Einrichtungen und die dort arbeitenden Wissenschaftlerinnen und Wissenschaftler profitieren. Gerade vor dem Hintergrund der Digitalisierung auch weiter Teile der Forschung in den Sozial- und Lebenswissenschaften ist ein gut ausgebauter Datenzugang essentiell für den Standort – auch um hervorragende ForscherInnen an heimische Institutionen zu rufen bzw. hier im Land zu halten. Zudem tut sich auch auf europäischer Ebene einiges, wie zum Beispiel die kürzlich veröffentlichte „Europäische Datenstrategie“ (Europäische Kommission

14 Insgesamt waren zum Zeitpunkt des Erscheinens des Artikels im Mai 2018 rund 273.000 Personen von ELGA abgemeldet. Das entspricht etwa 3% der im Jahresdurchschnitt 2018 anspruchsberechtigten Personen in der österreichischen Krankenversicherung. Siehe Hauptverband der österreichischen Sozialversicherungsträger, Tabelle 3.

2020), die als wesentlicher Baustein gesehen wird, die europäische Digitalindustrie wieder global wettbewerbsfähig zu machen. Die Vorteile, die sich für Evaluierungen – allgemein aber auch im FTI-Politik-Bereich – durch einen gut ausgebauten Datenzugang ergeben, müssen hier nicht weiter erläutert werden.

Aber eben nicht nur die österreichische Wissenschaft würde profitieren: Ein moderner und adäquater Zugang der Wissenschaft zu Register- und Statistikdaten ist die Voraussetzung für erfolgreiche evidenzbasierte Politik zum Nutzen der Allgemeinheit, der Steuerzahlerinnen und Steuerzahler ebenso wie des Wirtschaftsstandorts. Daher fordern wir: Zugang zu den Register- und Individualdaten für die Wissenschaft!

LITERATUR

Apuzzo, Matt, We Just Wanted to Talk E.U. Farm Policy. Why Was Someone Always Looking Over Our Shoulders?, *The New York Times*, 3. November 2019, <https://www.nytimes.com/2019/11/03/reader-center/eu-farm-subsidy-reporting.html>

Connelly, Roxanne, Dibbend, Chris, Gaylec, Vernon, Playford, Christopher J., The role of administrative data in the big data revolution in social science research, *Social Science Research*, Volume 59, September 2016, Pages 1-12, <https://doi.org/10.1016/j.ssresearch.2016.04.015>

dpa, Grüne warnen vor Plänen zu Gesundheitsdaten-Forschung, *Handelsblatt*, 3. November 2019, <https://www.handelsblatt.com/dpa/wirtschaft-handel-und-finanzen-gruene-warnen-vor-plaenen-zu-gesundheitsdaten-forschung/25184230.html>

Ecker, Brigitte et al., Österreichischer Forschungs und Technologiebericht 2019, Lagebericht gem. § 8 (1) FOG über die aus Bundesmitteln geförderte Forschung, Technologie und Innovation in Österreich, Bundesministerium für Bildung, Wissenschaft und Forschung, Bundesministerium für Verkehr, Innovation und Technologie, Bundesministerium für Digitalisierung und Wirtschaftsstandort, 2019, https://www.bmbwf.gv.at/dam/jcr:6c2909b9-d953-4575-8aed-50e705999f8a/Forschungs-_und_Technologiebericht_2019_dt_bf.pdf

ede/apa, Wer hat Zugriff auf Gesundheitsdaten?, *Wiener Zeitung*, 11. April 2018, <https://www.wienerzeitung.at/nachrichten/politik/oesterreich/958173-Wer-hat-Zugriff-auf-Gesundheitsdaten.html>

Europäische Kommission 2020, Mitteilung der Kommission. Eine europäische Datenstrategie, Brüssel 19.2.2020, https://ec.europa.eu/info/sites/info/files/communication-european-strategy-data-19feb2020_de.pdf Falk, Martin, Hölzl, Werner, Oberhofer, Harald, Die Bedeutung von unternehmensbezogenen Individualdaten für die empirische Wirtschaftsforschung und wirtschaftspolitische Beratung, Österreichisches Institut für Wirtschaftsforschung, WIFO Monatsberichte, 2015, Jahrgang 88 (11), S. 845-857, https://www.wifo.ac.at/jart/prj3/wifo/resources/person_dokument/person_dokument.jart?publikationsid=58522&mime_type=application/pdf

Gaulhofer, Karl, Wifo-Alarm: Statistik Austria blockiert Forschung, *Die Presse*, 09.07.2009, <https://www.diepresse.com/493556/wifo-alarm-statistik-austria-blockiert-forschung>

Hauptverband der österreichischen Sozialversicherungsträger, Versicherte Pensionen Renten, Jahresergebnisse 2018, <http://www.sozialversicherung.at/cdscontent/load?contentid=10008.555201&version=1548666221>

Kocher, Martin G., Badelt, Christoph, Manifest für eine evidenzbasierte Politik, *Der Standard*, 20. November 2019, <https://www.derstandard.at/story/2000111261182/manifest-fuer-eine-evidenzbasierte-politik>

Oberhofer, Harald, Schwarz, Gerhard, Strassnig, Michael, Registerforschung: Verwaltungs- und Statistikdaten für die Wissenschaft, *Mitteilungen der Vereinigung Österreichischer Bibliothekarinnen und Bibliothekare*, Bd. 72, Nr. 2 (2019), DOI: <https://doi.org/10.31263/voebm.v72i2.3154>, <https://journals.univie.ac.at/index.php/voebm/article/view/3154/2782>

Polt, Wolfgang, Streicher, Jürgen et al., Österreichischer Forschungs- und Technologiebericht 2018, Bericht der Bundesregierung an den Nationalrat gem. § 8 (2) FOG über die Lage und Bedürfnisse von Forschung, Technologie und Innovation in Österreich, Bundesministerium für Bildung, Wissenschaft und Forschung, Bundesministerium für Verkehr, Innovation und Technologie, Bundesministerium für Digitalisierung und Wirtschaftsstandort, 2018, https://www.bmvit.gv.at/dam/jcr:847e6be5-fb89-4b97-b90b-619d84a69538/ftb_2018.pdf

Stanzl, Eva, Forschung, Daten, der Schlüssel zum Wissen, *Wiener Zeitung*, 13. April 2018, https://www.wienerzeitung.at/nachrichten/wissen/mensch/958675-Daten-der-Schluessel-zu-Wissen.html?em_no_split=1

WZ online, APA, Tausende Abmeldungen von ELGA, *Wiener Zeitung*, 6. Mai 2018, <https://www.wienerzeitung.at/nachrichten/politik/oesterreich/963188-Tausende-Abmeldungen-von-ELGA.html>

AUTOREN

GERHARD SCHWARZ

WIFO - Österreichisches Institut für Wirtschaftsforschung

E: gerhard.schwarz@wifo.ac.at

THOMAS KÖNIG

IHS - Institut für höhere Studien und Wissenschaftliche Forschung

E: koenig@ihs.ac.at

HARALD OBERHOFER

WU - Wirtschaftsuniversität Wien

E: harald.oberhofer@wu.ac.at

MICHAEL STRASSNIG

WWTF - Wiener Wissenschafts-, Forschungs- und Technologiefonds

E: michael.strassnig@wwtf.at

KEYWORDS:

Registerforschung, Administrative Daten, Austrian Micro Data Center, Bundesstatistikgesetz, Statistik Austria

NEUES WAGEN – MIT DER SANDPIT-METHODE ZU NEUEN KONSTELLATIONEN & INTERDISZIPLINÄREN IDEEN

SANDPIT-ERFAHRUNGEN AUS SICHT EINER FÖRDERUNGSAGENTUR FÜR ANWENDUNGSORIENTIERTE FORSCHUNG UND INNOVATION

CHARLOTTE ALBER

DOI: 10.22163/fteval.2020.465

ABSTRACT

Mit der Pilot-Ausschreibung „Ideen Lab 4.0“ ist die Österreichische Forschungsförderungsgesellschaft (FFG) im Jahr 2018 neue Wege gegangen, um bereits in der **Phase der Ideen- und Antragsentwicklung neue ProjektpartnerInnen zusammen zu bringen** und die **Co-Entwicklung anwendungsorientierter Vorhaben** zu stimulieren. Die FFG eröffnete damit einen Experimentier- raum für eine neue Herangehensweise, um kooperative interdisziplinäre Projekte durch die **„Sandpit“-Methode** zu fördern. Eine externe Evaluierung wurde begleitend durchgeführt. Der vorliegende Artikel gibt, ergänzend zu der im Herbst 2019 fertig gestellten Evaluierung, Einblicke in die Programmentwicklung und Implementierung durch die FFG.

Neues wagen - das war die Leitlinie für die Ideen Lab 4.0-Ausschreibung im Jahr 2018 - sowohl für FörderungswerberInnen als auch für die FFG. Der Einsatz der Sandpit-Methode bedeutete für alle Beteiligten eine neue Herangehensweise, um interdisziplinäre, innovative Vorhaben zu generieren und neuartige Konsortien zu bilden. Interessierte mussten sich mit ihrem Lösungspotential, persönlicher Motivation und Unterstützungserklärung ihrer Organisation zu den Fragestellungen der Ausschreibung „Mensch 4.0? - Die Zukunft der Zusammenarbeit Mensch-Maschine“ bewerben¹. 30 Personen wurden von einer internationalen Jury hinsichtlich ihrer Innovationsfähigkeit, ihrer themenrelevanten Expertise, positiven Einstellung (gegenüber anderen Blickwinkeln, Disziplinen, Forschungsaspekten) und ihrem Lösungspotenzial ausgewählt. Die Gruppe, die zum Ideen Lab im September 2018 eingeladen wurde, war hoch interdisziplinär, mit unterschiedlichen Organisationshintergründen (Wirtschaft, Wissenschaft, Forschung, Gesellschaft) und Hierarchiestufen. Es erwartete sie eine fünftägige Veranstaltung, bei der sie sich in einem iterativen Prozess zu Teams zusammenfanden, Ideen entlang der Ausschreibungsfragen entwickelten, Kooperationen planten, Vorhaben prä-

sentierten, Peer-to-Peer Feedback erhielten und schließlich einen kurzen Förderungsantrag einreichten. Am Ende des Ideen Labs wurden von der vor Ort anwesenden Jury Empfehlungen zur Einreichung eines Vollantrags für Kooperationsprojekte (bis zu 1 Mio.€/Projekt) ausgesprochen.

MEHR ALS NUR BUNTE KLEBEZETTEL FÜR INTERDISZIPLINÄRE ZUSAMMENARBEIT

Sandpits mögen von außen betrachtet, aufgrund der Assoziation mit dem englischen Wort „Sandkasten“ und dem Einsatz unterschiedlicher Innovationsmethoden (und den damit verbundenen vielen bunten Klebezettel an den Wänden), verspielt wirken. Der durch die Agentur vorgegebene Rahmen (z.B. Art und Fokus der Förderungsprojekte, interdisziplinäre Fragestellung, Kooperationsformen) und die Ergebnisorientierung (Jurysitzung am Ende der Veranstaltung) machen sie jedoch zu mehr als nur unterhaltsamen Kreativ- und Networking-Workshops. Sie sind ein fordernder Prozess, der allen Beteiligten sehr viel Offenheit und Flexibilität abverlangt. In kurzer Zeit müssen eingelernte disziplinäre und organisationale Denkschemata abgelegt werden. Brücken für die Zusammenarbeit unterschiedlichster AkteurInnen werden gebaut und eine inhaltliche Basis sowie gemeinsame „Sprache“ für ein Forschungsprojekt gefunden. Es geht darum, sich auf Neues und andere Menschen einzulassen: Offene Kommunikation und Kooperation im gleichzeitigen Wettbewerb. Ein Sandpit rückt den einzelnen Menschen mit seinem bzw. ihrem Innovationspotential und Lösungsorientierung in den Mittelpunkt. Nicht die Person mit dem brillantesten Lebenslauf oder klingendem Organisationsnamen ist gefragt, sondern Menschen, denen

¹ Für die Ausschreibung, die vom Bundesministerium für Digitalisierung und Wirtschaftsstandort (BMDW) unterstützt wurde, standen insgesamt 3,5 Mio.€ aus Mitteln des Österreich-Fonds und der Nationalstiftung für Forschung, Technologie und Entwicklung zur Verfügung. Informationen zur Ideen Lab 4.0-Ausschreibung 2018 sind verfügbar unter der URL: <https://www.ffg.at/ideenlab/ausschreibung2018>

es gelingt, in kurzer Zeit ihre Fachexpertise im Austausch mit anderen Disziplinen und Persönlichkeiten zu einem neuen innovativen Vorhaben zu verbinden. Die durchgängige, mehrtägige Teilnahme kann Gruppendynamiken und individuelle Befindlichkeiten an die Oberfläche spülen. Gleichzeitig wird mit Hochdruck an dem Ziel eines erfolgreichen Antrags bzw. einer vorteilhaften Rolle für die eigene Organisation in einem potentiellen Konsortium gearbeitet. Dies kann für die Beteiligten neben dem Zeit-, auch Gruppen- und Erfolgsdruck bedeuten.

INNOVATIVE HERANGEHENSWEISE UND NEUE FUNKTIONEN FÜR DIE FÖRDERUNGSAGENTUR

Die mehrtägige Veranstaltung ist allerdings nur ein kleiner - für Externe aber besonders sichtbarer - Ausschnitt des Formats. Neben dem klassischen Programmdesign bzw. der Ausschreibungsarbeiten (z.B. Erstellung der Leitfäden, Mobilisierung, Vorbereitung der Auswahl Sitzungen) bedeutet es für die Förderungsagentur einen mehrmonatigen inhaltlichen Vorbereitungsprozess. Die Vorarbeiten können - abhängig von den Vorerfahrungen mit der Methode - zwischen sechs und zwölf Monaten dauern. In diesem Zeitraum werden von der Agentur die zentralen Stellschrauben (z.B. thematischer Fokus, Definition einer Fragestellung für die Ausschreibung und Zielgruppe, Umfang der notwendigen Interdisziplinarität, Art der zu fördernden Projekte, FachexpertInnen für die Jury) gestellt. Diese definieren den Rahmen dafür, was später im Ideen Lab entstehen kann. Beim mehrtägigen Ideen Lab tritt die Agentur in den Hintergrund und übernimmt erst am letzten Tag, mit der Abwicklung der Jurysitzung, wieder eine aktive Rolle. Die ersten Tage werden von externen ModeratorInnen und FachexpertInnen begleitet. Die FachexpertInnen fungieren am Anfang des Ideen Labs als MentorInnen und wechseln im späteren Verlauf in die Rolle der Jury.

Sandpits ermöglichen Agenturen neue Herangehensweisen zusätzlich zum „normalen Förderkanon“ auszuprobieren. So gelten sie zum einen als eine Alternative zum klassischen Peer-Review für die Auswahl von Projektanträgen, zum anderen stellen sie auch einen Rahmen für Kreativität und Interaktion zwischen den Antragstellenden und der Jury dar: *„Research proposals that undergo multiple rounds of review over a lengthy time period may suffer from ineffective interaction between researchers and reviewers, time delays and lingering uncertainty about when and whether projects will be funded. A sandpit model aims to address many of these issues. It brings together researchers, funders and reviewers to interactively discuss and revise proposals at a workshop, prior to making a final decision on which projects will be funded from a predetermined pot of money. Sandpits provide a forum for brainstorming sessions to foster creativity and generate substantive changes to proposals prior to their final submission. Since funding decisions are made at the end of a workshop, sandpits provide a finite timeframe for proposal review and revision. Appropriate selection of participants is crucial to the success of this approach, as well as effective facilitation/moderation by the funding administrator.“*^[1]

Sandpits sind aber nicht nur unter dem Aspekt eines neuen Auswahlverfahrens für eine Agentur interessant. Die FFG übernimmt dadurch Funktionen, die normalerweise die AkteurInnen selbst bei der Anbahnung neuer Kooperationen tragen müssen: Potentielle PartnerInnen werden im Zuge einer konkreten Antragsentwicklung für ein Projekt zusammengebracht und einigen sich mit neuen KonsortialpartnerInnen, dieses auch gleich umzusetzen. Hinzu kommt der, im Vergleich zu anderen Ausschreibungen, geringere Aufwand zur Einreichung eines Antrags für die FörderungswerberInnen. Bei einem Ideen Lab beträgt der zeitliche Aufwand für die Verfassung, Einreichung und Information über Förderungsempfehlung eines Projektantrages die Teilnahmezeit am Ideen Lab: wenige Tage. Damit werden Forschungskapazitäten, die sonst wegen der Akquise von Drittmitteln über Wochen blockiert sind, ohne zu wissen, ob man Chancen auf Erfolg hat, wieder für die Forschung frei.

NEULAND FÜR DIE ANWENDUNGSORIENTIERTE FORSCHUNGSFÖRDERUNG

Sandpits werden international vor allem im akademischen Bereich bzw. zur Förderung transformativer Grundlagenforschung eingesetzt. Es gibt wenig Erfahrungen, wie man die Methode einsetzt, um interdisziplinäre Projekte mit Anwendungspotential für die Wirtschaft zu fördern. Die FFG betrat dadurch Neuland [2]. Das Ideen Lab 4.0 wurde deshalb als Pilot-Ausschreibung umgesetzt, mit dem Ziel, Lernen für die Organisation zu ermöglichen. Für die Implementierung des Ideen Lab wurden von der FFG Hypothesen aufgestellt und extern begleitend evaluiert. Die Ausschreibung war damit selbst ein „Experimentierraum“, in dem Neues unter kontrollierten Bedingungen ausprobiert wurde. Durch die begleitende Evaluierung bekam die FFG rasch Hinweise zur Beurteilung des Erfolgs der Ausschreibung (anhand der Indikatoren im Programmdokument hinsichtlich der Erreichung der Programmziele). Andererseits lieferte sie Erkenntnisse, was die Sandpit-Methode für die Anforderungen der FFG leisten kann und was nicht. So standen für die FFG im Zuge der Implementierung folgende Fragen im Raum, wie: Trägt das Format dazu bei, neue Zielgruppen anzusprechen oder neue Projekt-Konstellationen zu generieren? Liefert der mehrtägige kreative Prozess des Ideen Labs neue Ideen? Entstehen durch den intensiven mehrtägigen Austausch neue Kontakte und weiterführende Forschungsaktivitäten? Wie wirkt das mehrstufige Verfahren auf die Auswahl interdisziplinärer Projekte? Da die Evaluierung die Erfahrungen aller am Sandpit-Beteiligten (TeilnehmerInnen, FachexpertInnen, ModeratorInnen, BeraterInnen, FFG) einholte, konnte die FFG diese im Jahr 2019 systematisch in das Design der zweiten Pilot-Ausschreibung Ideen Lab 4.0² einarbeiten.

AGENTUREN ZWISCHEN OFFENHEIT, INSTITUTIONELLEM RAHMEN UND RISIKO

Die FFG betrat das „Sandpit-Neuland“ aber nicht unvorbereitet. Der konkreten Programmentwicklung und Ideen Lab 4.0-Implementierung waren mehrjährige Vorarbeiten vorangegangen. In einer Impulsinitiative³ beschäftigte man sich mit dem Thema „Radikale Innovationen – Mehr Freiraum für innovative und risikobehaftete Forschung“ (2015-2016) und diskutierte dazu Aspekte wie Governance, Auswahlverfahren, Interdisziplinarität und neue Formate. Eine gleichlautende Studie und ein Arbeitskreis im Rahmen der Technologiesgespräche des Europäischen Forums Alpbach⁴ rundeten die Vorarbeiten ab. Ziel war es zu hinterfragen, wo man Freiräume für innovative Ideen im bestehenden Fördersystem schaffen kann. Als ein solcher Freiraum zeigte die Studie [3] die positiven Erfahrungen des Research Council of Norway (RCN) mit den „Idea Labs“ im Jahr 2016 auf. Der RCN setzt die Sandpit-Methode im Zuge seiner Strategie „Research for Innovation and Sustainability“ (2015-2020) als einen Baustein ein, um Prozesse neu aufzustellen und die Förderung innovativer Projekte zu ermöglichen. Der RCN orientierte sich bei seinen Idea Labs am Engineering & Physical Sciences Research Council (EPSRC). Der englische EPSRC hatte die Methode ursprünglich im Jahr 2003 zur Förderung transformativer, hoch interdisziplinärer Forschung entwickelt. Inzwischen werden Sandpits oder „Idea Labs“ weltweit veranstaltet. In den USA verwendete es die National Science Foundation [4] zur Förderung von „kreativen Ideen“, die im Standard-Schema nicht möglich wären. Universitäten stimulieren damit intern neue Kooperationsprojekte und interdisziplinäre Vernetzungen. In Österreich wurde das erste Ideas Lab von der Ludwig Boltzmann Gesellschaft (LBG) durchgeführt [5]. Im Frühjahr 2017 nutzte die LBG, als Forschungseinrichtung und Träger von Instituten, die Sandpit-Methode als neue Herangehensweise, um eigene Forschungsgruppen aus internationalen BewerberInnen zu formen. Das englischsprachige Ideas Lab war eingebettet in eine umfassende „Open Innovation in Science“ Initiative [6].

Das Beispiel des RCN und das wachsende Interesse⁵ für die Methode in Europa, zeigen, dass Agenturen - bedingt durch einen stetigen Wandel der Innovationssysteme - ihre Prozesse und Förderungen an die Bedürfnisse des Innovationsstandorts immer wieder anpassen müssen. Aktuell sind es beispielsweise neue Innovationsparadigmen, z.B. Open Innovation, die die Art und Weise wie Innovationsaktivitäten initiiert werden, ändern [7]. Auch die gesellschaftlichen Herausforderungen (Grand Challenges) erfordern aufgrund ihrer Komplexität vermehrt interdisziplinäre oder „radikale“ Lösungsansätze. Agenturen suchen dadurch nach Wegen, wie man die benötigten Lösungsansätze und das Wissen unterschiedlicher AkteurInnen wirkungsvoll stimulieren kann. Eine zentrale Rolle bei dieser Anpassung an den Wandel spielen für Agenturen

Auswahlverfahren. Denn: „*Gelingt es nicht in ausreichendem Maße jene Projekte, die für die Entwicklung eines Innovationssystems nützlich sind, von denjenigen, die das weniger sind, zu unterscheiden, werden die Ziele von Förderprogrammen nicht erreicht und potenziell öffentliche Gelder fehl alloziert.*“ [8] Wie ein Vergleich von 12 europäischen Innovationsagenturen zeigt, verwenden Agenturen deshalb auch eine Bandbreite von Auswahlverfahren, um ihrem jeweiligen Auftrag nachzukommen [9]. Diese Bandbreite, als auch der austarierte Einsatz der Verfahren, ergeben sich auch aus den unterschiedlichen institutionellen Rahmen, in welchen Innovationsagenturen agieren müssen.

LERNERFAHRUNGEN BEI DER IMPLEMENTIERUNG DES IDEEN LABS 4.0

Vor diesem Hintergrund gibt es für die FFG zentrale Lernerfahrungen bei der Programmentwicklung: Die Methode (also der gesamte Sandpit-Prozess) lässt viele Gestaltungsmöglichkeiten zu, z.B. im Hinblick auf die Zielgruppe, das Ausmaß der Interdisziplinarität, die Dauer des Ideen Lab, ob es nur eine erste Förderungsempfehlung zur Einreichung eines Vollertrags am Ende der Veranstaltung geben soll oder eine finale Förderungsentscheidung, die Art der Projekte und die Zusammensetzung der Konsortien. Man muss sich bewusst sein, dass jede Stellschraube, an der man dreht, die möglichen Ergebnisse (z.B. transformative Projekte generieren versus Stimulierung neuer Kooperationen) verändert. Das Design für ein Ideen Lab kann man deshalb auch nicht 1:1 nach dem Vorbild anderer Agenturen übernehmen. Es hängt stark von der konkreten Ideen Lab-Zielsetzung und dem institutionellen Rahmen der Agentur ab. Die FFG profitierte im Zuge der Pilot-Ausschreibung sehr von den Ideas Lab-Erfahrungen und der begleitenden Beratung der Ludwig Boltzmann Gesellschaft, konnte aber aufgrund der großen Unterschiede nur einen Teil der Erfahrungen auf das anwendungsorientierte FFG-Ideen Lab übertragen. Wichtig ist des Weiteren ein erprobtes Moderations-Team, das mit Sandpits vertraut ist. Wie bei allen anderen FFG-Förderungen unterliegt auch ein Ideen Lab dem Ziel eines effektiven und effizienten Einsatzes öffentlicher Mittel nach strengen Qualitätskriterien und verlangt damit transparente, möglichst standardisierte Abläufe. Für das Auswahlverfahren mussten einige Abläufe für die FFG neu entwickelt (z.B. ein vierseitiger Kurzantrag im elektronischen Einreichsystem für die Einreichung der Vorhaben am Ende des Ideen Lab) bzw. adaptiert werden (z.B. Detailliertheit der Kostendarstellung im Kurzantrag). Inhaltlich verfolgte die FFG mit der Pilot-Ausschreibung die gezielte Stimulierung von neuen interdisziplinären Kooperationen und die gemeinsame Umsetzung anwendungsorientierter Projekte. Die Konsortien konnten die im Ideen Lab erarbeiteten Vorhaben entweder als Sondierung, Kooperationsprojekt oder als F&E-Dienstleistung umsetzen. Diese verschiedenartigen Förder-

3 Geleitet vom BMDW (damals BMWFW) und von der FFG im Zuge des Programms w-FORTE begleitet. Innerhalb der vom österreichischen RatFTE unterstützten Initiative gab es auch organisationsübergreifende Arbeitsgruppen mit ExpertInnen aus Förderungsorganisationen, Ministerien und Stakeholdern.

4 Europäisches Forum Alpbach, Breakout Session 28.8.2016, Radikale Innovationen: Mehr Mut zum Risiko!, verfügbar unter der URL: <https://www.alpbach.org/de/session/tec16-breakout-session-07-radikale-innovationen-mehr-mut-zum-risiko/>

5 Im Mai 2019 fand ein internationaler „Sandpit-Learning Event“, organisiert von Netherlands Organisation for Scientific Research (NWO) statt, mit TeilnehmerInnen aus Österreich, Norwegen, Polen, England und Deutschland.

und Finanzierungsinstrumente⁶ ermöglichten der FFG und auch den Förderwerberinnen mit unterschiedlichen Förderhöhen, Zielsetzungen und Mindestvoraussetzungen zu experimentieren. Für die FörderwerberInnen bedeutete die Instrumenten-Vielfalt allerdings eine weitere Komplexität während des Ideen Labs. Die Förderung war auf Organisationen (je nach Organisationstyp mit unterschiedlichem finanziellen Eigenanteil) und nicht auf die Individualperson ausgelegt. Zielgruppe war aufgrund der geltenden Richtlinien die nationale FTI-Community, internationale Beteiligungen in den Konsortien waren in kleinem Ausmaß möglich. Während des mehrtägigen, deutschsprachigen Ideen Labs trat die FFG in einer neutralen, beobachtenden Rolle auf, was von den TeilnehmerInnen sehr positiv empfunden wurde.

Beim erstmaligen Einsatz der Sandpit-Methode für die anwendungsorientierte Forschung in Österreich stellten sich für die FFG auch Fragen hinsichtlich der konkreten Ausgestaltung der Ausschreibung, beispielsweise: Gibt es genügend experimentierfreudige InteressentInnen, die den zeitlichen Aufwand für die mehrtägige Veranstaltung und das Arbeiten an einer vorgegebenen Fragestellung in Kauf nehmen? Bekommen die TeilnehmerInnen von ihren Organisationen tatsächlich das Pouvoir, neue Vorhaben zusammen mit PartnerInnen, die man erst während der Veranstaltung kennenlernt, zur Förderung einzureichen? Werden innovative Projekte entstehen? Und welches Förderungsinstrument bietet genügend Flexibilität, aber auch Anreiz, für die sich im Ideen Lab bildenden Konsortien? Durch den Pilotcharakter der Ausschreibung und das hypothesengeleitete Vorgehen konnte mit verschiedenen Stellschrauben experimentiert und das Prozessrisiko für die Agentur kalkulierbar gehalten werden.

AUSWAHL UND VORBEREITUNG DER BETEILIGTEN ALS ZENTRALER ERFOLGSFAKTOR

Die Natur des Ideen Lab verlangt von allen Beteiligten einen hohen Grad an gegenseitigem Vertrauen, freiem Ideenaustausch, Offenheit für alternative Sichtweisen und interdisziplinäre Lösungsansätze sowie konstruktives Arbeiten im Team. In diesem Zusammenhang ist die Zulassung (Auswahl der BewerberInnen durch die Jury) und die Schaffung eines hierarchiefreien Raums (damit die Ideen möglichst offen ausgetauscht werden) während der Veranstaltung besonders wichtig. Die Auswahl der TeilnehmerInnen erfolgt deshalb nicht anhand von klassischen Exzellenzkriterien, sondern anhand von Fragen zur Person hinsichtlich Innovationsfähigkeit, themenrelevanter Expertise und positiver persönlicher Einstellung (bzw. besonderer Mehrwert für das Ideen Lab). Im Zuge der Zulassung muss mitbedacht werden, dass die für ein Ideen Lab ausgewählte Gruppe auch das formale Potential für die Bildung von späteren Konsortien mitbringt. So gibt es bei den FFG-Förderungsinstrumenten zum Beispiel Mindestkonsortien, die verpflichtend mindestens ein kleines oder mittleres Unternehmen vorschreiben. Für die Entwicklung innovativer anwendungsorientierter Projekte benötigt es auch technische

Umsetzungsexpertise im Konsortium. Neben einer gut ausbalancierten interdisziplinären Gruppe müssen somit auch diese Eigenschaften in einem anwendungsorientierten Ideen Lab vertreten sein. Ein interessantes Lernfeld aus der Pilot-Ausschreibung betrifft die Auswahl und Vorbereitung der FachexpertInnen (=Jury). Sie ist für die Agentur aufwändiger als bei anderen Formaten. Die Jurymitglieder müssen nicht nur fachlich für die interdisziplinäre Bewertungsrolle qualifiziert sein, sondern ebenso experimentierfreudig wie die TeilnehmerInnen. In den ersten Tagen des Ideen Labs sind die Jurymitglieder noch in einer MentorInnen-Rolle und sollten daher Coaching-Qualitäten besitzen. Für ein Jurymitglied ist der Aufwand für die Vorbereitung (diese beginnt zusammen mit der Agentur schon Monate vor dem Ideen Lab) und durchgängige Anwesenheit während eines Ideen Lab hoch. Hinzu kommt die persönliche Involvierung: Jurymitglieder sind für alle TeilnehmerInnen sichtbarer und interaktiver Teil des Prozesses.

ERGEBNISSE DER PILOT-AUSSCHREIBUNG IDEEN LAB 4.0 LAUT EVALUIERUNG

Betrachtet man die Ergebnisse der Pilot-Ausschreibung, zeigt sich, dass es der Aufwand wert ist, die Sandpit-Methode für die anwendungsorientierte Forschungsförderung einzusetzen. Laut Begleitevaluierung von Geyer/Good [2] ist „es der FFG insgesamt sehr gut gelungen, mit der Ausschreibung neue Akteure anzusprechen. Insgesamt wurden 112 Bewerbungen abgegeben, wobei 23% der Bewerber zuvor noch nicht zum Kundenkreis der FFG zählten.“ (S.13) Die 30 ausgewählten TeilnehmerInnen des Ideen Lab als auch die sechsköpfige Jury bestätigten übereinstimmend, dass „die Teamkonstellationen in den ausgearbeiteten und geförderten Vorhaben deutlich interdisziplinärer seien als in traditionellen FuE-Projekten, in die sie derzeit eingebunden sind oder in der Vergangenheit eingebunden waren“ (S.10). Die MentorInnen gaben auch an: „Es würden in den Projekten Akteure zusammenarbeiten, die schon in der Vergangenheit längst hätten zusammenarbeiten sollen“ (ebd.). Das Format würde im Vergleich zu konventionellen Ausschreibungen „neue Zielgruppen ansprechen und damit unübliche Konstellationen ermöglichen.“ (ebd.) Die Evaluierung kommt zum Schluss, dass „zwei der drei geförderten Projekte eine im Vergleich zum Durchschnitt der geförderten Projekte in anderen FFG-Programmen eine überdurchschnittlich hohe Bewertung des Innovationsgrads“ (S.13) aufweisen. Sehr erfreulich ist auch der hohe Anteil (47%) an teilnehmenden Frauen (und zwei der geförderten Projekte werden von Konsortialleiterinnen geführt). Die Jury war hinsichtlich des Verhältnis Männern und Frauen paritätisch besetzt. Das Vorurteil⁷, mehrtägige Sandpits führen zu wenig weiblicher Beteiligung, kann somit nicht bestätigt werden.

6 Die Instrumente der FFG sind verfügbar unter der URL: <https://www.ffg.at/instrumente>

7 Auch der EPSRC weist in seiner Sandpit-Reflexion auf die Wichtigkeit einer egalitär und divers aufgestellten Gruppe hin (und dass gerade die Geschlechterverteilung oft sehr ungleich ist). Verfügbar unter der URL: <https://epsrc.ukri.org/funding/applicationprocess/routes/network/ideas/whatisasandpit/sandpitpsychology3/>

RESÜMEE ZUR PILOT-AUSSCHREIBUNG

Sandpits sind eine geeignete Methode für die anwendungsorientierte Forschungsförderung, um verschiedene Disziplinen und heterogene Expertisen von unterschiedlichen AkteurInnen systematisch zusammen zu bringen. Ein Ideen Lab-Teilnehmer brachte seine Erfahrungen so auf den Punkt: „*Liebe FFG, Vielen Dank für alles, es war hart und mein Hirn raucht immer noch... aber ich glaube, daß derartige Formate, mit leichten Änderungen, eine wirklich sehr sinnvolle Ergänzung zum normalen Förderkanon sein können.*“⁸ Ideen Labs können durch eine entsprechende Formulierung der konkreten Ausschreibungs-Fragestellung⁹ auch sehr gut gesellschaftlich relevante Themenstellungen einbinden und dadurch ForscherInnen aus klassischen Technologiebereichen sowie aus den Sozial- und Geisteswissenschaften gleichermaßen ansprechen. Die durch die Sandpit-Literatur geschürten Erwartungen an das Ideen Lab, dass dadurch radikales Querdenken entsteht, haben sich nicht erfüllt. Es liegt die Vermutung nahe, dass dies auf die Adaptierung des, ursprünglich für die transformative bzw. Grundlagenforschung entwickelten, Formats zurück zu führen ist. Beim anwendungsorientierten FFG-Rahmen muss immer eine mögliche Umsetzung und Verwertung durch die beteiligten Unternehmen mitgedacht werden, was das radikale Querdenken einschränkt. Die FFG hat viele der in der Begleitevaluierung dokumentierten Lernerfahrungen (z.B. intensivere Vorbereitung aller Beteiligten, stärkerer Fokus in der Fragestellung auf die Anwendungsorientierung bzw. Beteiligung von Unternehmen, Ausschreibung von nur einem Förderungsinstrument, Höhe der Förderung) im Rahmen einer zweiten Ausschreibung umgesetzt. Im Dezember 2019 fand ein neues Ideen Lab (im Themenfeld „Künstliche Intelligenz und ethische Grundsätze“) statt. Dafür gab es wieder viele interessante und interdisziplinäre Bewerbungen. Das zweite FFG-Ideen Lab dauerte nur dreieinhalb Tage und es konnten am Ende Sondierungsprojekte (max. 200.000 €) eingereicht werden. Sondierungsprojekte ermöglichen ein rasches Austesten der innovativen Konzepte in einem überschaubaren zeitlichen und finanziellen Rahmen als Vorprojekt oder Anstoß für größere, innovative Vorhaben. Aufgrund des vereinfachten Antragsverfahrens wurde der vollständige Projektantrag bereits im Ideen Lab eingereicht und die Konsortien erhielten die Förderungsentscheidung noch rascher und ohne weiteres Antragschreiben. Von den 10 eingereichten Sondierungsprojekten wurden sechs vom Bewertungsgremium vor Ort zur Förderung empfohlen. Das Ideen Lab 4.0 ist damit ein Programm, in dem systematisch sowie effektiv, innovative interdisziplinäre Vorhaben in neuen Projektkonstellationen für den Innovationsstandort umgesetzt werden können.

LITERATURVERZEICHNIS

- [1] **S. Guthrie, B. Guerin, H. Wu, S. Ismail und S. Wooding** (2013): Alternatives to Peer Review in Research Project Funding: 2013 Update. Rand Corporation, Santa Monica, CA, p. 4. https://www.rand.org/pubs/research_reports/RR139.html
- [2] **A. Geyer und B. Good** (2019): Begleitevaluierung der Pilotausschreibung Ideen Lab 4.0 - Zusammenfassung der Ergebnisse. Inspire Research, Wien, S. 13-16, 2019. <https://repository.fteval.at/492/>
- [3] **K. Warta und T. Dudenbostel** (2016): Radikale Innovationen – Mehr Freiraum für innovative und risikobehaftete Forschung. Studie im Auftrag von BMWFW und RatFTE, Technopolis, Wien, 2016. <https://repository.fteval.at/191/>
- [4] **V. Gewin** (2012): Risky research: The sky's the limit. Nature, Bd. 487, Nr. July, p. 395–397, 19 July 2012. <https://www.nature.com/articles/nj7407-395a>
- [5] **K. Warta**, (2017): Lessons learned at the first Ideas Lab in Austria. Technopolis, Wien, 2017. <http://www.ideaslab.lbg.ac.at>
- [6] **LBG** (2017): Research Groups on Mental Health of Children & Adolescents, A novel approach to form Research Groups. Processdocumentation. Ludwig Boltzmann Gesellschaft, Vienna, 2017.
- [7] **BMVIT und BMWFW** (2017): Österreichischer Forschungs- und Technologiebericht 2017, Wien, 2017. <https://repository.fteval.at/318>
- [8] **P. Biegelbauer und T. Palfinger** (2016): Auswahlverfahren von Forschungsförderungsorganisationen im internationalen Vergleich. fteval Journal for Research and Technology Policy Evaluation, Bd. 41, p. 63, 2016.
- [9] **Biegelbauer, T. Palfinger und S. Mayer** (2017): How do innovation agencies evaluate and select projects? A comparison of 12 European agencies. fteval Journal for Research and Technology Policy Evaluation, Bd. 43, p. 21-23, 2017.

AUTORIN

CHARLOTTE ALBER

Österreichische Forschungsförderungsgesellschaft – FFG
(Strukturprogramme)

Sensengasse 1, 1090 Wien

E: charlotte.alber@ffg.at

KEYWORDS:

Begleitevaluierung, Sandpit, Ideen Lab, Interdisziplinarität, anwendungsorientierte Forschungsförderung

⁸ Email eines Ideen Lab-Teilnehmers an die FFG nach Absolvierung der fünftägigen Veranstaltung, 3.10.2018

⁹ Ein Beispiel hierfür ist die Fragestellung der 2. Pilot-Ausschreibung: „Wie können wir KI-Systeme bzw. deren Algorithmen unter Berücksichtigung ethischer Grundsätze möglichst vertrauensvoll gestalten, damit österreichische Unternehmen diese akzeptieren, ihre Potenziale erkennen und ausschöpfen?“

THE FUNDING SCHEME “IMPACT INNOVATION” – THE IDEA BEHIND AND THE RESULTS OF AN ASSESSMENT

PHILIPP AIGINGER-EVANGELISTI, BARBARA GLINSNER AND DOROTHEA STURN

DOI: 10.22163/fteval.2020.466

WHY “IMPACT INNOVATION”?

In February 2017 the Austrian Research Promotion Agency (FFG) started with a first call of the funding scheme “Impact Innovation”. It pursued three objectives:

- **Broadening the innovation base** through a broader understanding of innovation to include problem-based, non-technological innovation and innovation outside of traditional R&D-areas including social innovations, business model innovations and service innovation in diverse fields such as education, journalism and social inclusion;
- **Broadening access to FFG** funding for actors who have not been addressed by FFG instruments so far (e.g. start-ups, social enterprises, associations);
- More successful innovation projects through **an early involvement of relevant actors**, a structured and methodically guided approach along an innovation process, orientation towards problem solving and learning experiences on the problem in the course of project implementation.

Against the background of these objectives, the pilot programme aimed at innovation projects in their very early stages of the develop-

ment process, as Figure 1 illustrates. In contrast to traditional R&D-funding instruments an in-depth problem analysis, the ideation process and the idea selection formed an integral part of the projects to be funded. Therefore, the projects’ evaluation focused on the innovation process rather than on the research and/or market potential as it is usually done both by FFG and most European Innovation Agencies.

MOTIVATION FOR THE NEW PROGRAMME “IMPACT INNOVATION”

For the last decades Innovation Agencies used to fund innovations which were R&D- and technology-based. However, facing the societal and economic challenges and the likely transformational processes, efforts by research promotion agencies to support non-technological innovation could be noticed throughout Europe (e.g. the scheme “Feelings” from Tekes [now Business Finland] or the “Business Innovation Initiative” from Enterprise Ireland). In the FFG the first steps were taken with the introduction of the “service innovation initiative” and the “innovation voucher” scheme. But although both funding schemes were important

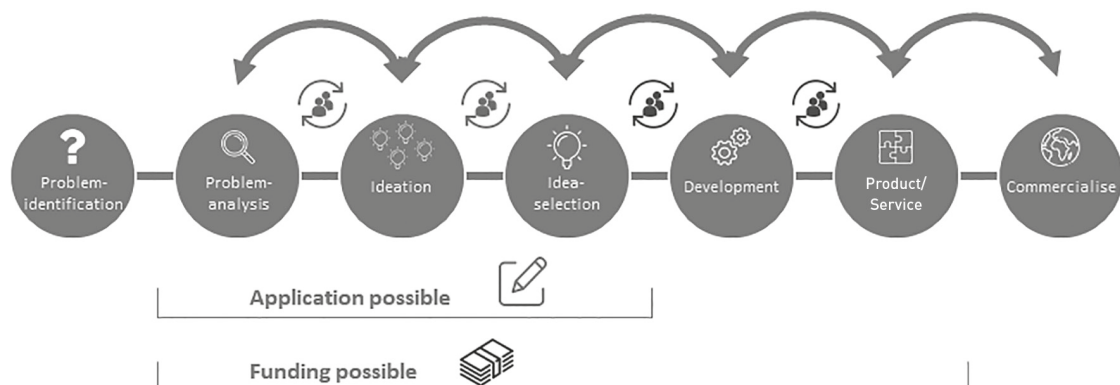


Figure 1: „Impact Innovation“: Interactive Development of innovative Ideas and Solutions, Source: FFG (<https://www.ffg.at/programm/impactinnovation>), own translation

milestones towards a new understanding of innovation by including non-technical innovations consciously, they were overall still rooted in the same narrow understanding of innovation as R&D driven.

So far, innovation was first and foremost conceptualised as the outcome of R&D activities that followed a linear process and resulted in new technological solutions (see Figure 2). This also becomes apparent when looking into the state aid rules for R&D aid, whose underlying rationale is a narrow concept of the market failure theorem.¹ The major argument in this context is the public good argument of R&D, which stipulates that new knowledge can be appropriated by competitors, thus hindering innovators from claiming 100% of the benefits² from their innovation while bearing all the risks connected with it. This assumed disproportion would lead to diverging private and social profits and non-optimal investments. Correcting market failure³ has traditionally been the core argument for the funding activities of Innovation Agencies.

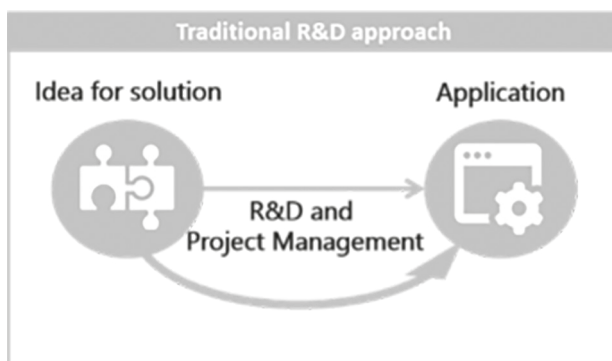


Figure 2: Traditional approach - Use of R&D to achieve the project goals
Source: FFG, internal use

However, this traditional focus of the Austrian Research Promotion Agency on R&D-based technological innovation failed to attract new forms of innovations, particularly non R&D based innovation. This shortcoming was also pointed out by some of FFG's network partners (e.g. chamber of commerce) and communities (e.g. start-ups) and became apparent when the FFG dealt with new target groups such as lean start-ups of the service economy. As a consequence, FFG started to rethink its approach by discussing with other Innovation Agencies of the TAFTIE⁴ network and by forming mainly internal working groups. Although the goal was to think openly and allow for new ideas, the persistence of the traditional approach based on the above mentioned narrow concept of market failure became obvious. As with the well-established FFG-Programmes, discussions about the sufficient level of innovation and risks of the projects were dominating the debate. Thus, updating the traditional

concept of innovation to include non-technological innovation as well as to enlarge the target group of FFG programmes to attract new, non-traditional innovators (such as social innovators) required a change in perspective as well as inspiration from outside the traditional innovation policy's ecosystem.

FFG therefore assigned a small innovation consulting company "WhatAventure" to get a clearer picture of innovation processes, modern innovation methodologies and the needs of non-technological innovators with regard to support structures and funding instruments. By including "unusual suspects" (for the FFG, such as social or cultural non-profit organisations) and Lead Users that the FFG considered innovative, but who didn't perform R&D (such as designers, social entrepreneurs) it became evident that the available funding instruments were not flexible enough to allow for iteration, learning and stakeholder involvement.

WHAT ABOUT THE MARKET FAILURE, WHERE ARE THE RISKS?

The market failure argument on information asymmetries, risk aversion and uncertainty can be mitigated through a more inclusive development process. In a paper about non-technical R&D the authors conclude that uncertainties are rooted in the user acceptance of new non-technical solutions (see Technopolis et al, 2016)⁵. Therefore, the approach to start with user's problems and involving them in the iterative development process as illustrated in Figure 1, seems very appropriate to overcome uncertainties. SMEs do not follow this path (enough) without public interventions, because of the constraints in resources (incl. finance, time, staff, networks) and a lack of qualification. This market failure argument, emphasizing the role of uncertainties, differs from the previous mentioned public good argument and substantiates the approach of insisting on a thorough innovation process.

The importance of the innovation process required a shift in assessing project proposals as well: Instead of assessing the level of innovation, technical risks to overcome (otherwise no real research is needed) and the market potential of the solution, it might be more important to assess the set-up of the innovation process and the future potential, if everything works as planned. So to say: instead of assessing the solution, start with the problem and how you want to proceed to solve it. The assumption is, that if there is a problem without sufficient solutions, any solution to this problem will have an impact. It seems more uncertain to assess market potential in two to three years (or even longer) than assessing the effects of problems now. The hypothesis here is „The bigger the effects of a problem, the bigger the impact of the solution “.

1 There is a long list of literature that evidences that appropriation is not an easy exercise, because of the cumulative, tacit nature of technical knowledge (e.g. Pavitt, 1984; 1987, 1999; Nelson & Winter 1982).
2 The idea of research as a public good goes back to Nelson (1959), Arrow (1962), and Buchanan (1968). For a taxonomy of various patterns of technical change see Pavitt (1984).
3 The market failure rationale includes also other than the 'public good of R&D' argument such as the information asymmetries and risk aversion argument, the divergence between private and social (investment) time preferences argument, the indivisibility and technical large-scale risks argument etc.
4 The European Network of Innovation Agencies, see <https://www.taftie.org>
5 The Oslo Manual 2018 (OECD 2018) as well as the European Innovation Scoreboard (European Commission 2019) also introduced innovations which are not R&D based.

PILOTING THE APPROACH

With this new approach to support also non-technical innovations, the notion of innovation became more diverse. However, there were still a lot of unknowns and uncertainties on how to design a programme for this purpose: How do these new types of projects look like? From which industries will the proposals come? What kind of organisations will apply? Is it the right approach at all? Since these are quite severe questions, FFG decided to use a design approach to develop and test each step. Partly with established network partners, partly with lead users and partly with newly established networks (social innovation community) small elements were tested (criteria, application, communication) and iterated before starting a funding call. Although over the course of 1.5 years FFG already learned a lot about new target groups, a lot of uncertainty about demand and understanding of the needs remained. This was why FFG decided to carry out a small pilot call. The mobilisation of the applicants was facilitated by the newly established networks, the Lead Users and multipliers that the FFG had identified.

The following table gives an overview of the key figures of the pilot call.

Project volume:	up to 150,000 €
Funding quota:	50%, up to 75,000 €
Duration:	up to 12 months
Number of submissions:	104
Accepted projects:	16 (= 15%)
Applicants with previous FFG experience:	37 (= 36%)
Target group:	SMEs including start-ups and enterprises in the course of formation, corporations and partnerships, associations, NPOs, individuals

Table 1: Key figures of the pilot call of “Impact Innovation”
Source: own elaboration

ASSESSING THE PILOT PROGRAMME

THE FFG ASSESSED THE PILOT PROGRAMME IN TWO STEPS:

First, the FFG assigned BlueLab with assessing the service design of the pilot call directly after the funding decisions to learn about the applicants’ understanding of the call, their motivations for applying (simplicity

and problem focus) and to better understand their innovation processes. Regarding the communication of the pilot programme it turned out that the reference model of an innovation process (see Figure 2) was simple but very effective to establish a common understanding, everyone involved could refer to. This quick evaluation enabled quick changes for the following call 2018.

Second, the Centre for Social Innovation (ZSI) was assigned with assessing whether the pilot call has met the programme’s objectives. ZSI’s approach comprised the following methods:

- First, a **descriptive analysis** of the applying organisations on the basis of the FFG database was compiled.
- Second, building on this analysis **semi-structured interviews** with successful and rejected applicants were carried out to assess the funding instrument at project-level. The aim was to understand how funded projects designed and implemented the innovation process, how they involved stakeholders and deployed innovation methods as demanded by the programme.
- Third, ZSI conducted a **focus group** with experts in the fields of innovation (policy), including start-ups, social entrepreneurship and crowd-funding experts as representatives of a broader innovation eco-system understanding to assess the impact of the new funding instrument at programme level and in the Austrian innovation system.

ATTRACTING NEW INNOVATORS

In the course of the Pilot call, 104 organisations or individuals submitted their proposals to the FFG of which only 16 were chosen for funding. The objective of attracting new applicants was well achieved. 65 percent of them were first-time applicants while only 17 applicants had already submitted more than 10 proposals during the last 20 years. Interestingly, experienced applicants were not more successful in obtaining funding than those who had no prior experience in submitting proposals to the FFG. “Impact Innovation” is only open to small firms and organisations; it was nevertheless surprising that the majority of organisations applying were micro-enterprises or sole entrepreneurs, only three had more than 50 employees. The thematic focus of the submitted proposals was categorised according to the **SDGs**⁶ by the FFG. In general, the submitted proposals could be attributed to eleven out of the 17 SDGs pointing to the heterogeneity of the problems the proposals wanted to tackle. Most submissions could be attributed to SDG 9 “Industry, Innovation and Infrastructure”, which is the traditional focus of FFG. However, only two of them were funded. Although it is difficult to draw conclusions regarding the approval rate per SDG given the low number of cases, it can be said that a third of the projects, which contributed in total to SDG 4 “Quality Education”, SDG 8 “Decent Work and Economic Growth” as well as SDG 15 “Life on Land” were chosen for funding. The funding instrument managed to attract a new group of applicants, that (i) mostly had no or little experience in applying to FFG, (ii) were SMEs (mostly micro-enterprises) and (iii) aimed at finding solutions for a variety of problems that are usually not within the scope of traditional funding instruments of FFG (see Table 2).

Project Focus	Applicant	Region
Social Inclusion of Elderly via Virtual Reality	SME (Start-Up)	Upper Austria
GEARing up Innovation in Diagnostics for Antibiotic Resistance	SME (Start-Up)	Vienna
Development of a Bayesian planning HR and ePlanning Tool in Health Services	SME (Start-Up)	Upper Austria
Virtual Reality as an HR Development Tool	SME (Start-Up)	Tyrol
Open Innovation Process Development for Scaling and Multiplying Impact - Impactful Preschool- and Parents' programmes	SME (non-profit)	Vienna
–Service Tool to Simplify Shopping and Delivery for Chinese Tourists in Europe (who mostly shop for their relatives and often don't get to experience the city they are visiting)	Sole Entrepreneur	Vienna
New Tool for Levelling out the Under-construction of Terraces to simplify and drastically shorten the installation process	SME	Upper Austria
Development of a community-based Learning Approach for Journalists to tackle new challenges in (traditional) media	Association	Vienna
Biomimicry Thinking Design Process as an innovation method for the craftsmen of the „Werkraum Bregenzerwald“	Association	Vorarlberg
Development of an Online-Platform to configurate and market funeral services	SME (Start-Up)	Vienna
Development of an AI Tool to assess the “easy to read” criteria of texts	SME (Social Enterprise)	Styria
Automatic Translation of Package Leaflets of Pharmaceutical Products in Sign Language	SME (Social Enterprise)	Vienna
Integrative Property Management that takes social and economic aspects into account	SME (Social Enterprise)	Styria
Development of Innovative Support Services for Social Entrepreneurs in Rural Areas	Intermediary Organisation	Vienna
Development of a tool to make use of local expert knowledge to assess natural hazards for traffic infrastructures	SME (Start-Up)	Tyrol
Development of an Orientation Guide for Feminine Hygiene Products	SME (Social Enterprise)	Vienna

Table 2: „Overview of the funded projects of the pilot call of “Impact Innovation”

Source: own elaboration

The interviews with applicants revealed that **the programme was attractive to them** for three reasons:

- First, in comparison to other funding schemes, the call was open to many forms of organisations and therefore constituted one of very few possibilities for associations or sole entrepreneurs to have their innovation activities funded.
- Second, as pointed out above, the broader definition of innovation enabled non-technological innovations such as service or process innovations, which are supported by only few other funding instruments in Austria.
- Third, as intended by the programme design, the interviewed applicants appreciated that the focus of “Impact Innovation” lies on the innovation process itself rather than on a pre-defined result. This focus gave applicants the freedom and the resources

to understand the complex issues they were facing, test their central assumptions, do prototyping and adapt their solutions if necessary, whereas traditional funding instruments often demand to follow the plan and solutions laid out in the application form.

One of the interviewees stated:

„Impact Innovation is a new idea, a new aspect which is treated only rarely. I think It is cool that the FFG is going to fund such projects.“
(own translation)

INNOVATION THROUGH NEW INNOVATIVE METHODS AND STAKEHOLDER ENGAGEMENT

In line with the focus on the innovation process, the programme made it compulsory for funded projects to apply innovation methods such as design thinking, hackathons or innovation workshops. Hence, all of the funded projects applied these methods, albeit to a different extent. On the one hand, those organisations or projects that were well-established and directed the funds towards developing a new activity, service or product in addition to their core business designed their innovation process quite carefully. They “innovated” in a specific setting, with dedicated time, resources and used creative, innovative or scientific methods. On the other hand, those organisations that were in their start-up phase handled the innovation process more flexible. Innovation methods as well as constant testing and adaption are an integral part of their tacit knowledge and hence their daily activities, but often not reflected. Furthermore, the funded projects all engaged their most important stakeholder (groups), mostly to obtain feedback and test central assumptions. However, none of the funded project fully integrated stakeholders into the innovation process. In conclusion, the assessment recommended a better explanation and trainings of innovation and co-creation methods for the applicants to ensure that they are well-understood and do fertilize the innovation process.

OUTCOMES AND IMPACTS AT PROJECT LEVEL

The pilot call of “Impact Innovation” can be regarded as **successful since every funded project produced valuable outcomes**. Many projects could test their central assumptions by engaging their key stakeholders. Moreover, the innovation process helped to improve the “solution-fit” of the developed process, service or product. Those projects that aimed at developing an app or a web-platform managed to create prototypes which are well advanced, but not yet market-ready. Interestingly, the engagement of stakeholders and the innovative solutions contributed to the improvement of existing products and services as well.

As described above, one of the programme objectives was to broaden the understanding of innovation to include non-technological innovation and innovation outside of traditional R&D-areas. The assessment identified three different kinds of innovations; two of them being non-technological:

- The first group of innovations can be labelled as innovative education/training and service concepts. Mostly by making knowledge more accessible through adapting it to the needs of their target groups those projects contributed to the empowerment of vulnerable groups, increased transparency for customers and encouraged actors in finding innovative ways of dealing with societal challenges themselves.
- The second group consists of projects developing digital services like platforms or apps to link stakeholders, improve services and the accessibility of information for their users.
- The third and smallest groups were technological innovations.

It is important to underline that the **largest innovation potential lied in the combination of the three forms of innovation**. With the exception of two projects, all developed innovative concepts and services either had a digital or a technological aspect. However, compared to traditional R&D the technological innovation here was minor and would have not been funded by traditional instruments. In the context of “Impact innovation” technological innovations constituted a means in delivering services to a (partly marginalized) target group.

Beyond those outcomes the funded projects also created **valuable impacts**. By engaging with stakeholders, the funded projects created vital links among those stakeholders and therefore contributed to a stronger integration of the field they are active in. Taking on the role as an innovation leader in their field, they could also increase their own visibility and reputation. In two cases the funding scheme “Impact Innovation” was even a vital factor in the constitution of the funded organisation. Furthermore, the interviews revealed that the applicants would not have been capable (mostly due to lack of resources) to find innovative solutions without obtaining funding by this programme. Therefore, it can be concluded that the programme filled a gap in the funding ecosystem.

Although all of the funded projects achieved great results, they are now faced with the **challenge of further developing their solutions** to make them market-ready. For well-established for-profit enterprises the further development of their innovations is a strategic decision of whether to invest own resources or not. Start-ups and social entrepreneurs, however, often lack the financial and time resources to carry their innovation process forward. Therefore, they need to find adequate follow-up funding opportunities. Given the scarcity of funding opportunities for social enterprises the lack of follow-up funding is threatening the innovation to be fully realised.

PROGRAMME’S AND ECOSYSTEM’S LEVEL

The “Impact Innovation” programme itself is an innovative initiative which fills a gap in the funding landscape. In the current innovation ecosystem, there is nearly no comparable funding scheme that works with such a broad understanding of innovation, which explicitly includes non-technological innovations, and starts at this early stage of the problem identification:

- It allows more time and resources to be invested in the innovation and development process. This enables stakeholders to become better involved and results in a targeted use of innovation methods.
- It enables the promotion of innovative activities for beneficiaries for whom no adequate funding programme exists in the field. Through the programme, some social enterprises have been able to implement innovation activities.
- With its broader concept of innovation, the programme occupies an almost unique position in Austria’s funding eco-system with a great potential for innovative projects.

Although the pilot programme has a perceptible impact on the innovation ecosystem, its explicit inclusion of social and educational enterprises also raised a number of issues that were discussed, particularly **within the focus group**. Consideration was for instance given to dividing the programme into two lines of intervention. A first line of intervention

could be aimed at social enterprises that provide innovative solutions for services in the public interest and do not operate in a market or quasi-markets, while a second line of intervention could promote market-oriented innovation with societal impacts.

Whatever the outcome of a decision on this issue will be, “Impact Innovation” is currently filling the gap in the creation of early-stage innovations, which is a unique selling point in the portfolio of innovation support schemes.

CONCLUSIONS AND RECOMMENDATIONS

Despite the consistently positive results that have been shown, the programme presents a number of challenges which require further reflection. These refer to the programme’s key objectives of addressing new types of innovation and new target groups. The assessment therefore ends with the following recommendations:

- Welcome new applicants more explicitly and develop specific formats for advising them
- Rethink a matching of FFG’s funding with crowd-funding
- Simplify formal requirements as far as possible
- Systematically give feedback to the applicants and projects
- Initiate exchange among projects
- “Mainstream” “Impact Innovation”: Integrate some of the programme’s characteristics into other funding schemes.
- Develop ways and methods to assess the projects’ impacts
- Communicate the programme and projects more clearly
- Address regional multipliers in all Austrian provinces

IMPLEMENTATION OF RECOMMENDATIONS

Some changes in the call 2019 directly refer to these recommendations, especially the ones on establishing networks between funded projects and the provision of tailor-made support for new applicants. In addition, the communication with regional multipliers outside Vienna was improved. In 2020, a pilot on Social Crowdfunding (matching public funding with crowd funding) will be launched. There is still some progress to be made to mainstream the “Impact Innovation” approach, but also to communicate the importance of thorough innovation processes – even beyond R&D – to applicants or potential applicants.

REFERENCES

Arrow, K. J. (1962): Economic Welfare and the Allocation of Resources for Invention. In Universities-National Bureau Committee for Economic Research, Committee on Economic Growth of the Social Science Research Council (ed): *The Rate and Direction of Inventive Activity: Economic and Social Factors*, Princeton University Press

Buchanan, J.M. (1968): *Demand and Supply of Public Goods*, Chicago: Rand McNally & Company.

European Commission (2019): European Innovation Scoreboard https://ec.europa.eu/growth/industry/innovation/facts-figures/scoreboards_en

OECD (2018): *Oslo Manual. Guidelines for Collecting, Reporting and Using Data on Innovation*, 4th Edition, Paris. <http://www.oecd.org/science/oslo-manual-2018-9789264304604-en.htm>

Nelson, R. R. (1959): The simple economics of basic scientific research. *Journal of Political Economy*, LXVII (3), June, 297-306.

Nelson, R.R. & Winter S.G. (1982): *An Evolutionary Theory of Economic Change*. Cambridge, Massachusetts: The Belknap Press of Harvard University Press.

Pavitt, K. (1984): Sectoral patterns of technical change: towards a taxonomy and a theory. *Research Policy* 13, 343–373.

Pavitt, K. (1987): Technological accumulation, diversification and organisation in UK companies, 1945-1983. DRC discussion paper, Science Policy Research Unit, University of Sussex.

Pavitt, K. (1999): *Technology, Management and Systems of Innovation*. Edward Elgar Publishing

Technopolis et al. (2016): *Ökonomische und verwaltungstechnische Grundlagen einer möglichen öffentlichen Förderung von nichttechnischen Innovationen* https://www.bmw.de/Redaktion/DE/Publikationen/Studien/studie-zu-nichttechnischen-innovationen.pdf?__blob=publicationFile&v=2

United Nations (2019): *The Sustainable Development Goals Report 2019* <https://unstats.un.org/sdgs/report/2019/The-Sustainable-Development-Goals-Report-2019.pdf>

AUTHORS

PHILIPP AIGINGER-EVANGELISTI

FFG – Austrian Research Promotion Agency
Sensengasse 1, 1090 Wien
E: Philipp.Aiginger@ffg.at

BARBARA GLINSNER

ZSI – Centre for Social Innovation
Linke Wienzeile 246, 1150 Vienna
E: glinsner@zsi.at

DOROTHEA STURN

ZSI – Centre for Social Innovation
Linke Wienzeile 246, 1150 Vienna
E: sturn@zsi.at

KEYWORDS:

non-technological innovation, innovation processes, social innovators, stakeholders’ involvement

FUNDING RESEARCH DATA INFRASTRUCTURES: FUNDING CRITERIA IN GRANT PEER REVIEW

RICHARD HEIDLER

DOI: 10.22163/fteval.2020.467

ABSTRACT

Researchers increasingly recognise that data-sharing, improved data management and the availability of research data infrastructures serves the effectiveness of the science system. Funding agencies support this cause and react to it. The German Research Foundation (DFG) designed the funding programme “Information Infrastructures for Research Data” to address this need. The programme was evaluated in 2019 in respect to the long-term success and sustainability of the funded projects. The full study contains funding statistics, a descriptive content analyses of review panel protocols and qualitative case studies. For the content analysis, an in-house team at the DFG coded the content of review group protocols for 158 programme applications. An expert commission accompanied the original evaluation and offered recommendations for the further development of the programme. This article builds on this study and adds a statistical analysis of the programme specific funding decision based on the original data of the content analysis of review panel protocols. Which criteria are decisive for funding selection? The multivariate analysis of the funding decision models the outcome of the grant peer review under consideration of the arguments and assessments made by the review panels. The results show that the assessed quality of the research plan has the strongest effect on the funding decision. This is in line with other studies on peer review criteria. More programme-specific is the conclusion that the scientific demand for the planned research data infrastructures has a marked impact on the funding decision, too. An unexpected result is the low importance of the institutional environment for the final funding decision.

INTRODUCTION

In an increasingly collaborative and data-intensive science system, researchers emphasise the necessity for data sharing infrastructures (Tenopir et al. 2011). As a result, science policy makers and funding agencies pay attention to the prerequisites of data sharing and data practices that foster the accessibility, discovery, re-use and preservation of research data. Even though many researchers agree on the advantages of data sharing (Piwowar et al. 2007, Tenopir et al. 2011), there are still many reasons for researchers to withhold data, and sometimes

also practical constraints exist (Borgmann 2012). In addition, research cultures concerning data sharing practices vary observably for different research fields (Tenopir et al. 2011).

The German Research Foundation started a call for applications in 2010, and set up an ongoing funding programme from 2013 onwards where researchers can submit a funding proposal at any time. An additional second call in 2015 called “research data in practice” focused more on the practical utilisation of research data infrastructures as “proof of concept” in concrete research projects. The aim of the whole funding programme “Information Infrastructures for Research Data” is to improve data sharing and to support the development of tools and infrastructures in order to facilitate the distribution and access to quality assured research data. *“The funding programme aims to support science and the humanities in drawing up and implementing specific and needs-oriented requirements for future structures, or structures undergoing further development, for the improved handling of research data and research data repositories.”*¹ Applicants for the programme are researchers cooperating with scientific information-infrastructure institutions such as libraries, archives, computing and media centres, or information science experts.

Researchers applied for a funding volume of €135m, of which the DFG approved €48m for funding. The average funding volume ranges from €489,000 in the 2010 call to €668,000 in the 2015 call and €780,000 in the ongoing funding programme.

To evaluate the long-term success and sustainability of the funded projects, the DFG studied the implementation of the programme in depth. The full study contains funding statistics, a descriptive analysis of the funding criteria and qualitative case studies, carried out by the Institute for Innovation and Technology (iit) at VDI/VDE Innovation + Technik GmbH, Berlin.

A central question of the evaluation study was to analyse according to which criteria the review panels of the DFG choose applications for funding. For this purpose, an in-house team at the DFG analysed the content of review group protocols for 158 programme applications quantitatively. The full report contains the case studies and the descriptive analysis of the review protocols including a detailed code scheme with examples for every code (DFG 2019a).

An expert commission formed by members of the Committee for Scientific Libraries and Information Systems (AWBI) and further information management experts accompanied a programme evaluation. It also of-

ferred recommendations for the further development of the programme (DFG 2019b).

For this article, we add a multivariate analysis of the funding decision by modelling the outcome of the grant peer review under consideration of the arguments and assessments of the review panel to the descriptive analysis of the original study. The multivariate analysis serves to identify which review criteria were decisive for the funding decision and to weight their contribution to the final funding decision while controlling for contextual factors (e.g. field of research, type of application etc.) that may also affect the success rates.

RESEARCH QUESTION

The peer review and decision finding process in this programme aims to ensure the allocation of the funding to the best-suited proposals and applicants. At least one expert on the very research field and another for the infrastructure side (information science related expertise) write a review of each proposal. Subsequently, a review panel discusses the proposals based on the written reviews, assesses the quality of the proposals and proposes which projects should receive funding.

Reviews contain arguments for or against the funding of a proposal, measured against the aims of the respective funding programme and thus reveal how the grant review system is working. Nevertheless, content analysis of review documents is scarce and the substance of grant peer review is seldom analysed. Thus, the results of our analysis may lead to a deeper understanding of the peer review process (Reinhart 2010, 317). Although there is already some research on the practices and funding criteria used for classic grant peer reviews (Hartmann and Neidhardt 1990, Langfeldt 2001, Reinhart 2009, 2010), we know quite little about the adaption of reviewing practices within funding programs like the one that is at the centre of this study. Funding criteria for the promotion of information infrastructures of research data need to include additional and different criteria compared to classical grant peer review such as

- the consideration of standards and interoperability in the project proposal,
- the demand for data-sharing activities in the field,
- the acceptance by the scientific community,
- the information science quality
- a sustainable involvement of infrastructures sites.

“Classic” criteria for the assessment of research grants like the originality of the research question are much less important.

In the conclusions, we will discuss the question to which extent the review criteria differ from classical research grant peer review and how they fit the goals of the programme “Information Infrastructures for Research Data”. Furthermore, we discuss in the conclusions, to what extent the content analysis contributed to the further development of the programme.

DATA

The basis of the analysis are all applications submitted since 2010 until July 2017. Since not all 213 proposals were dealt within the review panel, the number of cases as a starting point for the content analysis is

reduced to 158 proposals (Table 1). Out of these proposals, 64 were accepted (including one or more renewal applications per project), belonging to 52 funded projects. This excludes applications, which have been decided by written procedure, by other means (e.g. a subcommittee), and a total of 24 applications which were recommended for rejection without discussion in peer the review panel because of the negative written reviews.

Table 1: Number of proposals by programme variant

	Applications discussed in review panel		
	Total	Accepted	Rejected
Call “Information Infrastructure for Research Data” 2010	63	28	35
Call “Research Data in Practice” 2015	46	11	35
Ongoing Funding Programme 2010 to 2017	49	25	24
Total	158	64	94

The analysis is facilitated by the fact that the minutes of the review panel and the structure of these meetings are very uniform. The review panel minutes represent a synthesis of the review statements from the written reviews previously obtained and the panel reviewer's arguments, which is why they are well suited for a content analysis of the review criteria.

METHODS

A quantitative content analysis of the review panel protocols of 158 applications serves to systematically identify and differentiate the diverse criteria for the funding decision in this programme. Hartmann and Neidhardt (1990) used a similar procedure, however focusing on written reviews, not on review panel protocols. They examined 639 reviews of 242 applications in the German Research Foundation. Reinhart (2009) performed a content analysis of 212 written reviews for 68 randomly selected applications from biology and medicine at the Swiss National Science Foundation (SNSF).

The use of the review panel protocols instead of the written reviews for the content analysis may have consequences for the model of the review decision. Reinhart (2010, p. 318) argues that panel discussion may increase the pressure to produce a consistent decision and thus lead to an increased importance of “social” factors in comparison to “scientific” or “epistemic” aspects: “panels seem to be especially susceptible to negotiation and mediation among the panellists; therefore, an agreement on fair decision-making is a more pressing need than shared epistemic standards. It could be argued that peer review without panels is superior when it comes to enforcing epistemic standards.” To study these aspects in depth, a separate and additional coding of the written reviews would have been necessary. However contradictory statements from written reviews are taken into account indirectly, since in such case review panel protocols contain contradictory statements in respect to the assessed dimension. Usually in such cases the protocols first report statements from the written review, and then give contradictory evaluation of the review panel. Overall 33 percent of the review panel protocols contain a

contradictory statement in (mostly) one dimension, the final model will test, whether such contradictory statements have an independent effect on the funding decision.

The development of a code plan (Table 2) containing the review statements was an iterative process based on the review panel protocols. Various steps ensured the validity and reliability of the coding and the code scheme. In a first step, the expert commission accompanying the programme evaluation worked through the review panel protocols to provide information on a possible code scheme. Review statements typically contain judgement statements; the coding focuses on such statements that contain either a negative or a positive assessment of a certain aspect of the proposal. The iterative process led to 94 individual subcodes that can be aggregated with regard to their content as negative or positive judgement in 24 distinct code dimensions. A total of 1,858 unique subcodes from the 94 subcode types were assigned in the respective applications, corresponding to 11.8 subcodes per application. Subcodes are only counted once as positive or once as negative in each of the 24 different code dimensions.

Based on this 24 more detailed dimensions we identify four main review criteria that were brought up as topics in the review panel protocols:

1. The scope of the scientific demand for the suggested research data infrastructure,
2. the quality of the concrete implementation concept (i.e. the research plan),
3. the expertise of the proposing scientist and
4. the adequateness of the institutional environment.²

To give an example: A subcode of the scientific demand is the question whether the applicants have performed a demand analysis. A positive valence can be given to statements that indicate that the demand analysis has been performed, like: *“The demands analysis was convincingly performed due to the presentation of various already ongoing activities and preliminary work.”* A negative valence would be assigned to statements that remarks that such an analysis is missing, e.g.: *“In addition, statements on the target group and a comprehensible needs analysis is missing.”* A negative valence was also coded if a similar existing infrastructure has not been taken into account, like in this statement: *“Finally, the first rapporteur points out that a number of similar systems already exist, which is not referred to in the application.”*

We give a second example for the review criterion “research plan”. A sub-code of this concept is the adequateness of the amount of money requested (“suggested costs”). Statements of adequate requested resources like: *“Time, personnel and resources are all calculated in a comprehensible manner”* are coded with a positive valence. Statements of too low (*“The second reviewer considered the requested resources, in particular the personnel capacities, to be undersized and assumed that the applicants underestimated the task”*), too high (*“In addition, the comparatively high budget applied for and the staffing level were not plausibly explained and comprehensibly justified”*) or unclear (*“The expenditure is also not calculable or comprehensible for lack of quantity information”*) resource plans were coded as negative statements.

Table 2: Code plan

4 Review criteria	24 Dimension	Positive subcode (+)	Negative subcode (-)
Scientific Demand	General	Very high	Unclear
		Existent/high/relevant	Weak
		Will increase	
		Increases efficiency	
		Generally necessary	
		Improvement	
		Otherwise data will be lost	
	Acceptance in scientific community	Well integrated	Acceptance uncertain
			Not involved
	Usability und user	High and well defined	Unclear
			To narrow
			Not reflected
	Internationality/supra-regional	International relevance	Regional relevance
		Supra-regional relevance	
	Environment and needs analysis	Performed	Missing
		Similar project not considered	

² The assignment of the individual codes to the code dimensions is described in Appendix A of the original study with examples for every subcode and definitions for all 24 code dimensions.

Research plan	General	Well elaborated	Inconsistent
		Coherent and plausible	Unclear
		Potential	To abstract
			Bad
	Fit to the program	Very good	Unclear
		Good	Bad
	Previous work	Outstanding	Not enough
		Succesful/good	
	Suggested costs	Appropriate	To high
			Unclear
			To low
	Originality	Yes	Unclear
			No
	Theoretical relevance	Very high	Weak/unclear
		High	
	Methods	Very good	Unclear
		Good	Weak/outdated
	Data/Licenses	Metadata well defined	Metadata unclear
		Data protection addressed	Meta data not well defined
		Licenses clear	Data protection issues
		Data quality high	Licenses unclear
		Quality control existent	Data quality unclear
			Data quality bad
		Quality control not existent	
	Standards	Considered	Not considered/unclear
	Feasibility	Realistic	Unclear if feasible
			Not feasible
Project goals	Interesting	Not well defined	
	Well defined		
Interoperability	Ensured	Not ensured	
Subsequent use	Ensured	Unclear	
		Not ensured	
Expertise	General	Excellent	Partially
		High reputation	Not good
	Information science/IT	(Very) good	Not covered
	Subject specific	Excellent	Subfield not covered
High reputation			
Institutional environment	General/subject specific	High reputation	Important facility missing
		Good collaboration	Not sufficient
	Infrastructure	High reputation	Cooperation problematic or missing
	Sustainability	Guaranteed	Unclear
			Missing concept or facility
			Not guaranteed

The coding was performed with the qualitative analysis software MAXQDA, the further quantitative analysis of the data was then carried out with the statistical software STATA. The reliability of the codes was systematically tested. Based on a random sample, 30 review panel protocols were coded a second time by the same person. This procedure helped to further specify the code definitions and to reduce ambiguities in the code plan. Finally, a second person coded 30 randomly selected review panel protocols, to measure the inter-coder reliability. Cronbach's alpha was calculated at the level of subcodes and the resulting value of 0.85 can be classified as "good".

ANALYSIS

The central goal of the analysis is to model the effect of the positive or negative assessments of the four different main criteria addressed by the reviewers on the funding decision. Since the dependent variable of the model is binary (funding or no funding), we applied a logistic regression.

The model includes several structural control variables that may influence the funding chances. The two calls and the ongoing funding programme are differentiated, since for the second call the funding rate was lower due to the high number of proposals. Furthermore, the 16 renewal proposals (which occur only in the ongoing program), have a higher success rate and this information is therefore included as control variable.

As additional multinomial control variable, the scientific discipline of the proposal is included. If the proposal is affiliated to more than one of the four main scientific disciplines, the proposal is regarded as multidisciplinary (16 percent of all proposals). The highest demand for the programme can be found in the Humanities and Social Sciences and the Life Science (both 35 percent), less proposals belong to the Natural Sciences (8 percent) or the Engineering Sciences (5 percent).

Where necessary and reasonable, infrastructural facilities like libraries or computing centres should be involved in the planned project. Therefore, the participation of an infrastructure facility is included in the

model. This is the case for one fifth of the proposals. As last additional control variable which may have an effect on the funding decision we include the amount of the funding requested (avg. €613,000 for the 158 analysed cases) as continuous variable in the model.

To analyse the effect of the assessment of the main review criteria on the funding probability, for all four dimensions, respective indicators are constructed. The four indicators root in the 24 dimensions described in the code plan in Table 2. Every dimension can be mentioned positively or negatively if at least one positive or one negative subcode occurs. Dimensions can also be coded positively and negatively at the same time, if at least one negative and one positive subcode was given. However, such contradiction in the assessment of the same dimension occur rather seldom, as Figure 1 depicts. The figure shows the number of dimensions with only a positive, negative or with a contradictory assessment, differentiated between approved and rejected applications. The analysis shows that funded and declined applications differ markedly in their valuation: While for approved applications a positive assessment is reported on average for five dimensions and an average of 1.3 dimensions is decided negatively, the ratio is almost reversed for rejected applications (4.3 negative versus 1.6 positive).

In addition, it is noticeable that the review panel uses for single applications only a fraction of the 24 maximum possible assessment dimensions, which the code plan identified across all applications. On the one hand, this is because some assessment dimensions occur rather rarely overall. In addition, if no positive, negative or contradictory rating, for example for "sustainability", was recorded this does not necessarily mean that sustainability is (not) given, but only that the matter of sustainability of the proposed project was not addressed in the peer review group protocols. This may also be due, for example, that this aspect was already treated as unproblematic in the written reports.

The complete study gives a more detailed and qualitative interpretation of the different dimensions. For the purpose of this article, we focus on the quantitative analysis of the four main review criteria. For the model of the funding decision, the results of the coding are further ag-

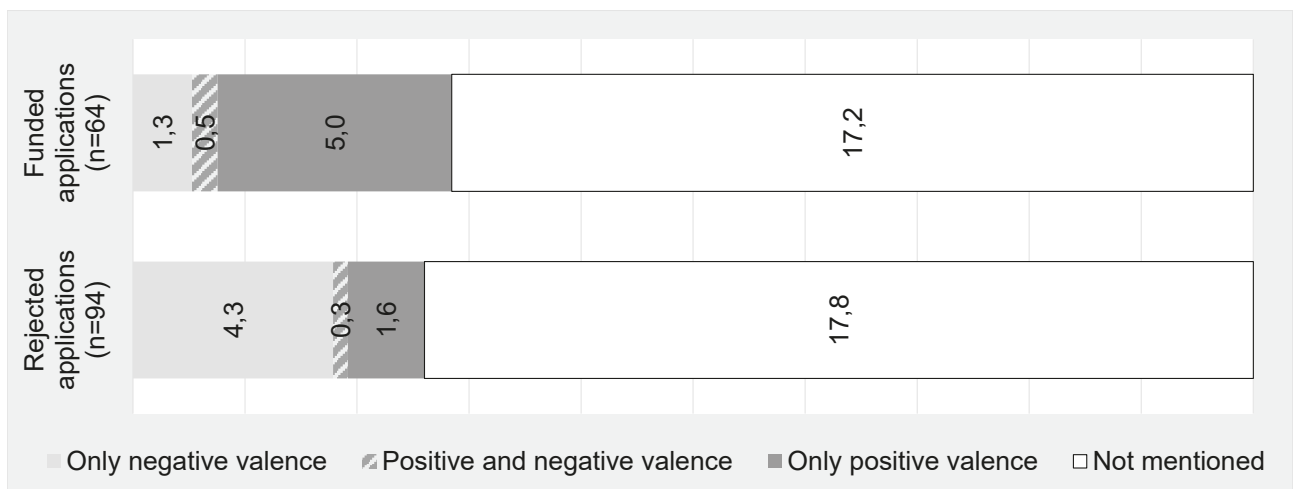


Figure 1: Average number and valence of rated dimensions (total 24)

gregated to generate a continuous indicator. We computed for every application for all four review criteria a value for a scale variable, containing the cumulated valences in each of the four main dimension (Figure 2).

Some review criteria comprise more dimensions than others: the expertise and the institutional environment contain three dimensions, the scientific demand five dimensions, and the research plan 13 dimensions. Inside each review criterion each dimension can add or subtract the value one (or both at the same time for contradictory statements) to the respective cumulated indicator value. The review criterion “research plan” thus has a theoretical range from -13 (if all dimensions would have at least one negative subcode and no positive subcode) to +13 (vice versa). The range for the review criterion “scientific demand” goes from -5

to +5 and for “expertise” and “institutional environment” from -3 to +3. To account for these differences in variation for the final model all four scale variables were treated as continuous variables and z-standardised, to make the coefficients more comparable.³ We interpret the resulting review criteria indicators as a measurement of the net valence in the number of dimensions with positive or negative assessments as noted in the review panel protocols for the respective topic. A negative value indicates that more dimensions have been rated negative than positive, a positive value vice versa. Figure 2 shows the distribution of the height of the indicators values for the four review criteria for accepted and rejected applications.

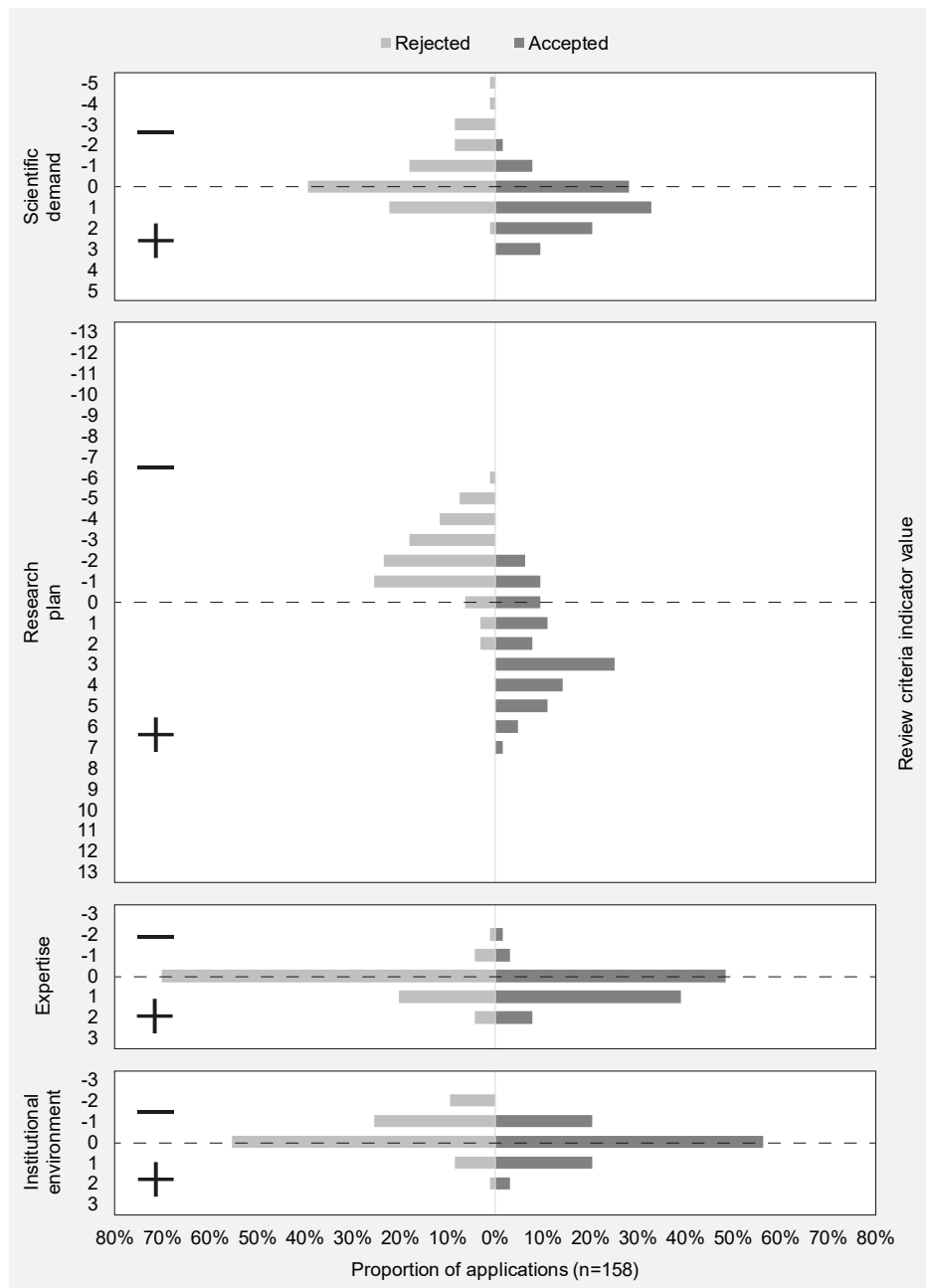


Figure 2: Distribution of the four review criteria indicators by funding decision

3 All other scale variables in the model are z-standardized, too.

The descriptive analysis (Figure 2) reveals that funded applications reach more often a positive indicator level especially for the “research plan” (75% versus 6%) and the “scientific demand” (63% versus 23%), with less pronounced differences for “expertise” (47% versus 25%) and “institutional environment” (23% versus 10%).

In how far do the indicators measure independent aspects of the assessed quality of the applications? A correlation analyses of the four review criteria indicators shows, that, although all four are positively

correlated, the correlations are rather low, which supports the assertion that the four identified criteria can be regarded as distinct aspects of the applications which are reviewed independently from each other. Scientific demand and the research plan correlate highest with 0.38, the correlation of the indicator values is the weakest between the scientific demand and the expertise (see Table 3).

Table 3: Correlation matrix of review criteria indicator values, all applications (n=158)

Indicator values	Scientific demand	Research plan	Expertise	Institutional environment
Scientific demand	1			
Research plan	0.384	1		
Expertise	0.140	0.294	1	
Institutional environment	0.192	0.204	0.297	1

The results of the final model can be found in the coefficient plot (Figure 3) and the model in Table 5. The overall model has a rather high explanatory power with a pseudo r^2 of 0.73. The predictive power of the model originates mainly from the four review criteria indicators, since the

pseudo r^2 of a model including only the control variables is considerably lower with 0.12 (not depicted).

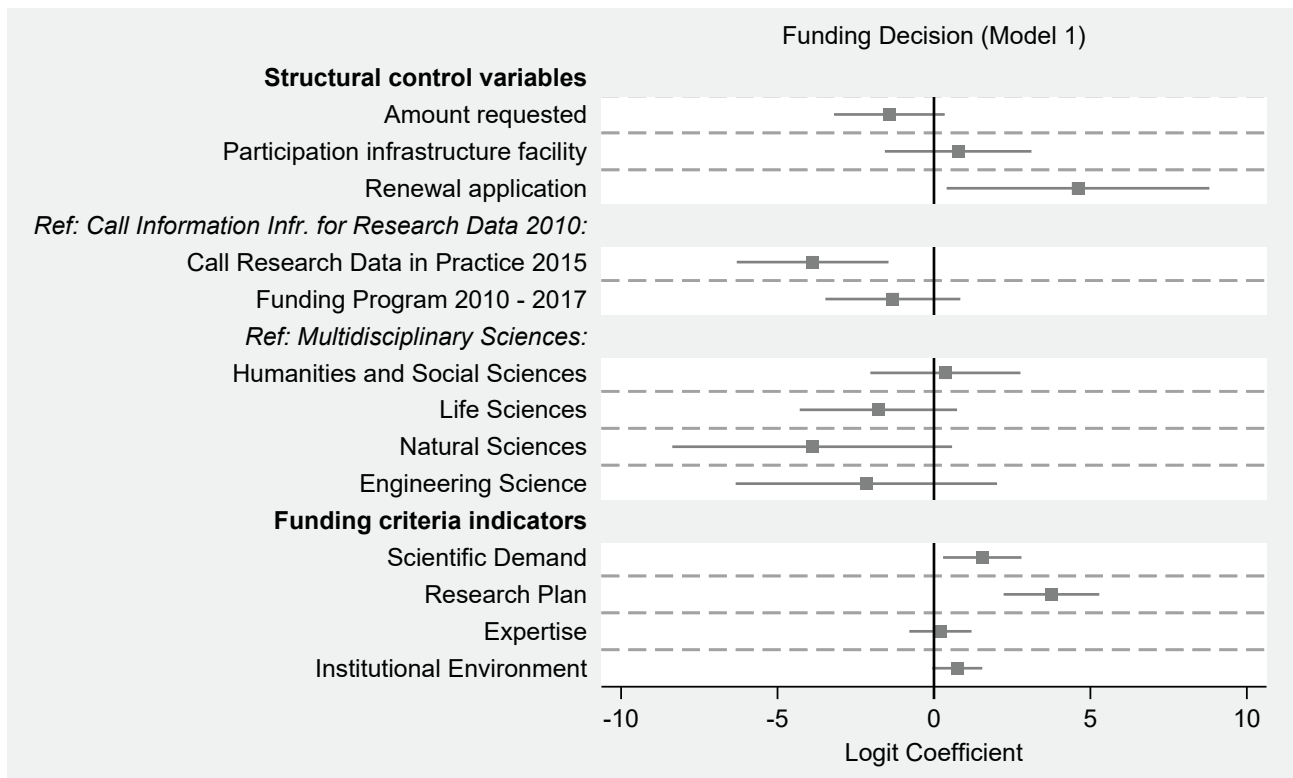


Figure 3: Logit coefficient plot model 1 funding decision (compare table 5)

In respect to the control variables, only the submission as a renewal proposal has a significant positive effect, while being part of the 2015 call "Research Data in Practice" has a significant negative effect on the funding probability. Higher amounts of money being requested has a slightly negative effect on funding chances, whereas the participation of an infrastructure facility increases chances, however both effects are not significant (see Figure 3). Although applications for larger funding volumes would be possible, the programme seems to be more attractive for projects with a smaller scale in the life science and social sciences and less for highly international and collaborative data and facility intensive fields like Climate Science, Particle Physics, Space Science or Bioinformatics.

Applications from natural and engineering sciences have a negative logit coefficient compared to multidisciplinary proposals; however, differences between disciplines are not significant. The effects of the four

possible, but is not presented here for the sake of the simplicity.

review criteria (under control of other variables) reveals relevant differences in detail: While in all four dimensions more positive statements have a positive predictive coefficient for the funding probability, the size and significance of the coefficient is varying.

Since all variables are z-standardized, both the logit coefficients and the p-values can indicate the varying importance of the four funding criteria. To make the effect of changes in the net number of positive and negative dimensions in each of the four review criteria more visible we additionally computed funding probabilities for three exemplary indicator values, holding all other variables constant at the means. Table 4 gives the funding probabilities for applications with one net positive dimension, a zero sum, or a net value of one negative dimension for each of the respective criteria (Table 4). The computation of predicted funding probabilities for higher or lower indicator values than +1 and -1 would be

Table 4: Predictive probabilities for funding at the means, for three different values for the four review criterion indicators

Review criterion indicator value	Funding probability at the means			
	Scientific demand	Research plan	Expertise	Institutional environment
1	51.3%	68.2%	32.6%	54.2%
0	26.3%	37.0%	26.4%	31.8%
-1	10.8%	13.9%	21.1%	15.6%

The strongest predictive value can be found for the "research plan". Holding all other effects constant, the marginal effect at the means gives a funding probability of 68% for applications with a net value of one positive assessed dimension, compared with a clearly lower funding probability of 14% for applications with one net negative statement in this topic. The coefficient is highly significant with a p-value below 1%. Slightly weaker is the effect of the assessment of the scientific demand for the application by the review panel, while still being significant with a p-value below 5%. The comparison of the predictive probabilities shows

a funding chance reaching from 51% for an indicator value of +1 to 11% for an indicator value of -1. As reflected in the funding probabilities, which differ only at a small scale for indicator values from +1 to -1, the assessment of the expertise has only a minor and non-significant effect on the funding decision. Finally, the institutional environment has a positive effect on funding probability; however, the coefficient is slightly above the significance limit of 5%. Although the difference in the predicted funding probabilities is still visible, reaching from 54% to 16%.

Table 5: Logistic regression model of the funding decision

	Funding decision (Model 1)			Funding decision (Model 2)		
	Coef.	Std. err.	p-value	Coef.	Std. err.	p-value
Structural control variables:						
Amount requested	-1.425	0.90	0.11	-1.402	0.9	0.12
Participation of infrastructure facility	0.773	1.19	0.52	0.842	1.19	0.48
Renewal application	4.604	2.14	0.03*	4.403	2.29	0.05+
Reference category: Call 2010						
Call "Research Data in Practice" 2015	-3.878	1.24	0.00**	-3.827	1.24	0.00**
Ongoing Funding Programme 2010-2017	-1.314	1.1	0.23	-1.361	1.1	0.22

	Funding decision (Model 1)			Funding decision (Model 2)		
	Coef.	Std. err.	p-value	Coef.	Std. err.	p-value
Reference category: Multidisciplinary Sciences						
Humanities and Social Sciences	0.365	1.22	0.77	0.379	1.24	0.76
Life Sciences	-1.776	1.28	0.17	-1.701	1.31	0.20
Natural Sciences	-3.892	2.28	0.09+	-3.807	2.3	0.10+
Engineering Sciences	-2.161	2.13	0.31	-2.036	2.15	0.34
Funding criteria indicators:						
Scientific demand	1.544	0.64	0.02*	1.513	0.65	0.02*
Research plan	3.755	0.78	0.00***	3.783	0.79	0.00***
Expertise	0.207	0.51	0.68	0.161	0.52	0.76
Institutional environment	0.745	0.41	0.07+	0.744	0.40	0.07+
Number of contradictory dimension codes				0.148	0.34	0.66
Constant	0.835	1.22	0.49	0.774	1.24	0.53
	N = 158, Pseudo R ² = 0.7323, LR chi ² = 156.19			N=158, Pseudo R ² = 0.7332, LR chi ² = 156.39		
	+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001					

The question in how far the funding decision would differ if only written reviews were taken into account and in how far contradictory assessments from the written reviews and the review panel affect the decision remains open. Although there is no separate analysis of the written reviews, we can at least indirectly test this assertion. Out of the 158 cases, 45 cases have contradictory statement in one dimension, five cases in two dimensions, two cases in three and one cases in four dimensions. This count variable is included as additional z-standardised control in the described model (compare model 2 in Table 5). We find no significant positive or negative effect of the number of review dimensions with contradictory statements on the funding probability. Furthermore, the coefficients' height and significance of the four review criteria indicators remain unchanged between model 1 and model 2. Nevertheless, additional research is necessary to study the interplay of written reviews and review panels.

CONCLUSIONS AND OUTLOOK

The analysis showed that it is possible to model the funding probability of applications based on a content analysis of review panel protocols while controlling for structural parameters, and to weight the importance of different review criteria. The importance of the research plan for the funding decision is in line with the results of Reinhart (2009) and Hartman (1990) for standard research grants.

The low relevance of the assessment of the expertise of the applicants for the funding decision goes in hand with an overall reluctance to raise this topic by reviewers found in other studies. This corresponds to the observation by Reinhart and Hartmann "that reviewers display goodwill and civility towards the applicant, an attitude that he sees rooted in the norm of not speaking negatively about colleagues" (Neidhardt, 1988:119). An alternative explanation would be that the expertise of the applicants is difficult to discern.

What seems to be more specific for this programme is the importance of the "scientific demand" for the funding decision. This aspect includes on the one hand the supply side, by assessing whether applicants did an environment and needs analysis, and defined well the users and usability of the envisioned data infrastructure. On the other hand, this dimension also looks at the demand side, where the reviewers assess the scope of the demand of such an infrastructure and the potential acceptance in the scientific community.

The multivariate model shows that the institutional environment is less decisive than expected based on the stated programme goals. This is in line with the findings of the evaluation study and the recommendations of the expert group accompanying the evaluation. Although the expert group was not in possession of the presented multivariate analysis, their conclusions in respect to the review process, based on the descriptive evaluation study and the case studies, are covered by the results presented here. They stated that the review process with written reviews in combination with panel discussions works well and should continue. Less under reviewer scrutiny than expected was the assessment of the institutional environment, including the long-term (financial) sustainability of the infrastructure. The expert group concluded - based on the complete study including the qualitative case studies - that the reviewers too seldom addressed this aspect and there is a tension between the innovativeness of a proposal and the long-term sustainability. They recommended: "the criteria for sustainability of the project results should be defined more precisely and revised. This requires a clearer identification of an expected sustainable operation of the scientific data infrastructure. Sustainability is to be evaluated in relation to the innovativeness of the projects" (DFG 2019b, 4). To reach this goal, they suggest splitting the programme in three segments differentiating "mature major projects" and community building "initial major projects" from innovative high-risk projects. In the last segment, demands for sustainability are weakened and replaced by the minimal expectation of further usage and integration of tools in existing infrastructures, while innovativeness is more central.

ACKNOWLEDGEMENTS

I am grateful to Anke Reinhardt for her valuable feedback and constructive suggestions for both the analysis and the article and Stefan Winkler-Nees for adding to the qualitative text analysis and accompanying the whole project.

REFERENCES

Borgman, C. L. (2012): The conundrum of sharing research data. *Advances in information science*, 63, 1059-1078.

DFG (2019a): Bewertung des Förderprogramms „Informationsinfrastrukturen für Forschungsdaten“. doi.org/10.5281/zenodo.2636516

DFG (2019b): Weiterentwicklung des Förderprogramms „Informationsinfrastrukturen für Forschungsdaten“. doi.org/10.5281/zenodo.2650866

Langfeldt, L. (2001): The Decision-making constraints and processes of grant peer review, and their effects on the review outcome. *Social Studies of Science*, 31, 820–841.

Piwowar, H. A., Day, R. S., Fridsma, D. B. (2007): Sharing Detailed Research Data Is Associated with Increased Citation Rate. *PLoS ONE* 2: e308.

Reinhardt, M. (2009): Peer review of grant applications in biology and medicine. Reliability, fairness, and validity. *Scientometrics*, 81, 789-809.

Reinhardt, M. (2010): Peer review practices: a content analysis of external reviews in science funding. *Research Evaluation*, 19, 317-331.

Hartmann, I. & Neidhardt, F. (1990): Peer review at the Deutsche Forschungsgemeinschaft. *Scientometrics*, 19, 419-425.

Tenopir, C., Allard, S., Douglass, K., Aydinoglu, A. U., Wu, L, Read, E., et al. (2011): Data Sharing by Scientists: Practices and Perceptions. *PLoS ONE* 6: e21101.

AUTHOR

RICHARD HEIDLER

Information Management, German Research Foundation
Kennedyallee 40, Bonn, 53175 (Germany)
E: richard.heidler@dfg.de

KEYWORDS:

Peer Review, Research Data, Data Sharing, Grant Peer Review, Review Criteria, Deutsche Forschungsgemeinschaft, Programme Evaluation

ANALYSIS OF THE INNOSUISSE START-UP TRAINING IN A THEORY OF PLANNED BEHAVIOR FRAMEWORK¹

ELISABETH NINDL AND PETER KAUFMANN

DOI: 10.22163/fteval.2020.468

ABSTRACT

The present paper analyzes the Innosuisse start-up training program in an empirically well-tested socio-psychological model, the Theory of Planned Behavior (TPB). The aim is to answer the question, whether the training modules (aiming at high-tech start-ups) enhance the participants' intention to found a business, and which determinants contribute to a change in this intention. Data has been collected in three consecutive surveys for all courses offered in the years 2016 and 2017. The estimation results indicate that those who feel empowered show a higher intention, while others realize during the course that a business foundation is not what they really want (at that point of time). The largest impact on the change in the intentions to found a business comes from the perceived behavioral control of the participants, i.e. the knowledge about running a business; to a somewhat lesser extent, but

still significantly positive, is the perceived social support. Overall, the Innosuisse start-up training helps student and university members to come to a more informed decision in favor or against a potential business formation and to develop their skills and knowledge.

BACKGROUND

The Innosuisse start-up training aims at supporting students and university members who are interested in setting up knowledge-intensive and technology-based firms. The general program objectives comprise (1) the fostering of an environment that allows young knowledge-intensive firms to develop successfully, (2) the promotion of an entrepreneurial mind-set at Swiss universities and (3) the conduct of practice oriented and interconnected training modules for students, university members and young (student) entrepreneurs (Kägi et al., 2018).

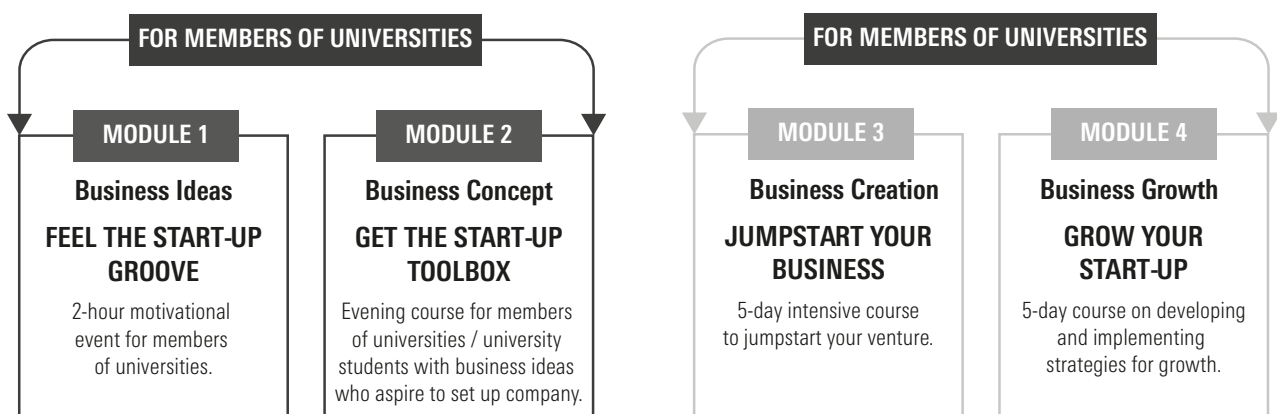


Figure 1: Start-up training modules 1-4

Source: <https://www.innosuisse.ch/inno/en/home/start-and-grow-your-business/start-up-training.html> (16.10.2019)

¹ This paper draws on the evaluation report by Kägi, et al. (2018). We thank our collaborators from BSS Research and Consulting for their contributions during the study, and our client from Innosuisse for constructive discussions. During the evaluation period, the name of the program was 'CTI Entrepreneurship Training Program'.

The four modules are:

- **Module 1:** Awareness raising for entrepreneurship in the knowledge- and technology-based sector; focus on members of universities and universities of applied science: students, doctoral students, postdocs, research fellows; duration two hours
- **Module 2:** Provision of skills for founding a business, identification of promising business ideas, encouragement of the participants to start their own business, promotion of an entrepreneurial spirit; focus on members of universities and universities of applied science: students, doctoral students, postdocs and research fellows with an entrepreneurial idea in the knowledge- and technology-based sector; duration ten evening sessions à four hours (one semester)
- **Module 3:** Assessment of the market potential of business ideas, identifying success factors and common pitfalls, planning and realizing business ideas; focus on business founders (shortly before or after setting up their company); duration five days spread over several weeks
- **Module 4:** Moving from the first clients and partners to a growing company; focus on founders and heads of knowledge- and technology-based companies already operating in the market; duration five days spread over several weeks

On behalf of the Swiss Commission for Technology and Innovation (KTI, since 2018 Innosuisse), BSS Research and Consulting and the Austrian Institute of SME Research were commissioned to perform a program evaluation of the start-up training program “CTI Entrepreneurship” (today “Innosuisse start-up”). The evaluation was carried out in the years 2016 to 2018 and used a mix of qualitative and quantitative methods. It covers the years 2016 and 2017 and analyzes the structure of the participants, the conception, realization and effect of the modules two to four. The training modules were offered by five leading institutions (universities, universities of applied sciences, business parks) and 29 partner institutions in four Swiss regions (Zürich East, Region Central, Region West and Tessin) (Kägi et al., 2018).²

In this paper, we focus on the direct outcome of the courses, which is the immediate behavioral effect on the target group, i.e. whether the course participants intend to take the next steps towards entrepreneurship. We could not observe during this study whether the course participants really set-up a business (impact), because this requires often a longer time horizon. The time from the first business idea to its successful implementation may even take several years and an impact analysis could thus be a fruitful task for future research.

In the following, we introduce the theoretical framework and its incorporation into the program’s intervention logic. Second, we give an overview of the data. In a third step, we describe the regressions and present selected estimation results including its interpretation.

THE THEORY OF PLANNED BEHAVIOR

The empirical evaluation of the effect of the start-up training on the participants builds on the Theory of Planned Behavior (TPB), an empirically well-tested model from socio-psychology (Ajzen, 1991). According to the model, the intention to perform a certain behavior in a specific context can be explained and predicted by the attitude towards this behavior, the individual’s subjective norm (the influence by important others) and its perceived behavioral control (PBC).

Attitude refers to the individual subjective perception, sensation and assessment of a certain behavior and its desirability. In contrast to this inner appraisal, subjective norm accounts for social factors such as the perceived social support, disapprobation or pressure for a specific behavior from the closer social environment (friends, family or peers). Perceived behavioral control, as the third component, refers to the perception of the ease or difficulty to perform the behavior of interest. It describes one’s confidence in the ability to perform a behavior and can be influenced by the individual’s choice of activities and the preparation for an activity, the expended effort, the thought patterns and emotional reactions. On these grounds, the elements of perceived behavioral control can be stimulated through education and training.

As is shown in Figure 2, the three factors attitude, subjective norm and PBC can encourage or discourage the intention to perform a behavior. In addition, the TPB assumes that the PBC not only affects a certain behavior via the channel of intentions, but also exhibits a direct effect on the behavior, which is in our context the founding of a business. Due to this twofold effect, perceived behavioral control is regarded as the most important component of the TPB.

The following Figure shows a part of the overall intervention logic (output to impact) of the program, in which the TPB is embedded. Nationwide coverage, the participation of the target group (i.e. the intended outreach of the program) and the quality of the training are the direct outputs of the program implementation. These outputs should translate into outcomes in the target (treatment) groups. We focus in this article solely on the (measurable) changes in attitude, subjective norm, perceived behavioral control and intentions. The empirical results from other studies show that intentions constitute a reliable indicator for future entrepreneurial behavior (e.g. Heuer and Kolvereid, 2013, Liñán and Fayolle, 2015, Miralles et al., 2016).

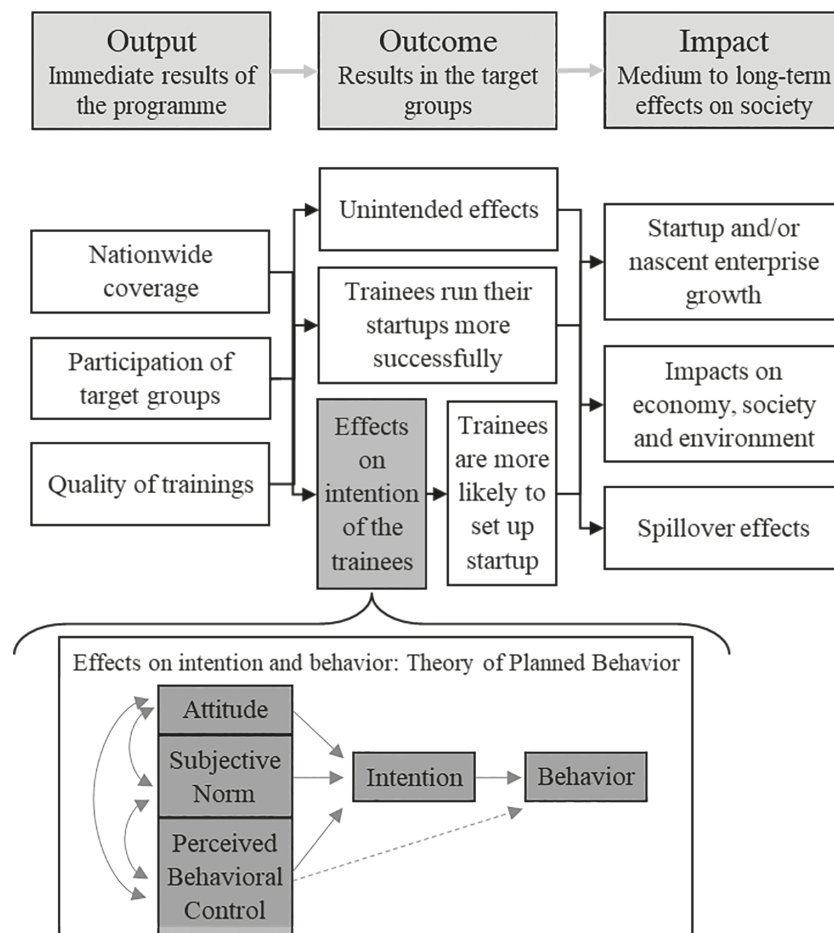


Figure 2: Embedding the TPB to capture the course outcomes in the intervention logic
Illustration by Kägi et al. (2018); TPB: Ajzen (1991)

DATA AND DESCRIPTIVE STATISTICS

Following the conceptualization of the study, the data was collected by BSS Research and Consultancy in specifically designed surveys. For modules two, three and four, the course participants were asked to answer three surveys: the first at the beginning of each course (QI), the second asks the same items after the end of the course (QII) and the third survey six months after completion of the training program (QIII). Five questions relate to the participants' attitude towards creating an own business, three items measure subjective norm, another three items focus on perceived behavioral control. Intention was measured using four items. Additional control variables are gender, age, the parents' entrepreneurial background, previous work experience in a small or newly founded firm and educational attainment.

In this article, we present only the results for module 3; the results for module 2 are qualitatively and quantitatively similar. Due to the low number of participants in module 4, it cannot be analyzed in the structural equation framework.

In the evaluation period 2016 and 2017, all 678 participants of module 3 were invited via email to complete the survey. The first questionnaire at the beginning of the course was answered by 451 participants (67%), the second survey by 402 (60%) and by 201 (30%) in the third wave six months after the course. Due to missing answers, another 30% of the observations were discarded from the sample. Table 1 shows the descriptive statistics of the assessment in QI for the final sample of 347 participants answering the first and the second survey, and 158 individuals answering all three surveys. The values for both samples are very similar, so that we can rule out that our results are driven by the group composition of the respondents.

	QI ∩ QII	in %	QI ∩ QII ∩ QIII	in %
Intention				
Choice: own business vs. employee	6.16		6.24	
Goal: creating a business one day	5.99		5.98	
Readiness: to do anything to become an entrepreneur	5.30		5.28	
Likelihood: Incorporate/register a company in the next 2 years	5.96		6.08	
Attitude				
Disappointing – Rewarding	6.36		6.35	
Bad for me – Good for me	6.22		6.22	
Worthless – Worthwhile	6.46		6.44	
Negative – Positive	6.43		6.46	
Harmful – Helpful	6.18		6.13	
Subjective Norm				
Support of your family	5.60		5.72	
Support of your friends	5.51		5.58	
Support of your colleagues	5.23		5.28	
Support of your entrepreneurial peers	5.59		5.72	
Perceived Behavioral Control				
Decision Making: Decision making under risk and uncertainty	5.34		5.33	
Probability of Success: Probability of succeeding with an own firm	5.05		5.10	
Time Sacrifice: Sacrifice the necessary time and energy to the project	5.97		6.12	
Background/Control variables				
Female	80	23.1%	38	24.1%
Age	34.9		35.3	
Entrepreneurial parents	122	35.2%	58	36.7%
Previous work experience	180	51.9%	85	53.8%
Student Uni	11	3.2%	8	5.1%
Student A.S.	19	5.5%	10	6.3%
Graduate Uni	61	17.6%	40	25.3%
Graduate A.S.	41	11.8%	23	14.6%
Foreign education	22	6.3%	14	8.9%
Other education	24	6.9%	17	10.8%
Education missing	169	48.7%	46	29.1%
N	347		158	

Table 1: Descriptive statistics, Module 3 (means)

Source: Kägi et al. (2018), own calculations

A.S. ...University of Applied Sciences. All variables (excluding background variables) ranked on a 7-point Likert scale with higher values corresponding to a better assessment. The table displays the means of the variables and, where possible, the shares of total.

In general, the participants' intentions to become an entrepreneur are very high – especially the choice for starting a business – while there is less agreement on the readiness to do anything to become an entrepreneur. Also, the attitude towards entrepreneurship is very high and only little variation can be found between the five measured items. The perception of the support from the closer social environment is slightly lower and more heterogeneous. Also the PBC items are assessed quite heterogeneously – while the participants are highly willing to sacrifice the necessary time and energy, they rate their abilities of decision making under risk and the probability of succeeding with his or her own firm lower (but still at a high level).

Between the two surveys QI and QIII (see Table 2), we observe significant and quantitatively large decreases in the intention items on the likelihood of creating a business in the next two years and the trainees' readiness to do everything to become an entrepreneur. For the items of perceived behavioral control, the trainees report a significant increase in their perceived probability of succeeding with an own firm and at the same time a significant and large decline in the assessment of the willingness to sacrifice the necessary time and energy for running a business. These findings already indicate that the training may have caused a somewhat more realistic assessment of entrepreneurship: trainees learned about the required time and effort and that they might not be willing or able to make these contributions; those who still wish to become an entrepreneur now feel a higher probability of succeeding (see also Oosterbeek et al., 2010). This is supported by a significant increase in the valuation that running an own business is worthwhile.

	QI	QIII	Change
Intention			
Choice	6.16	6.17	0.01
Goal	6.00	6.02	0.02*
Readiness	5.38	5.23	-0.15**
Likelihood	6.00	5.36	-0.64***
Attitude			
Disappointing – Rewarding	6.35	6.35	0.00
Bad for me – Good for me	6.22	6.24	0.02
Worthless – Worthwhile	6.44	6.54	0.11**
Negative – Positive	6.46	6.41	-0.05
Harmful – Helpful	6.13	6.12	-0.01
Subjective Norm			
Support of your family	5.72	5.48	-0.24
Support of your friends	5.58	5.44	-0.13
Support of your colleagues	5.28	5.03	-0.25
Support of your entrepreneurial peers	5.72	5.50	-0.22
Perceived Behavioral Control			
Decision Making:	5.33	5.45	0.11
Probability Success	5.10	5.29	0.20*
Time Sacrifice	6.12	5.72	-0.40*

Table 2: Change in the TBP-items between surveys QI and QIII

Source: Kägi et al. (2018), own calculations. N=158; significant differences for p-values<0.1: *...p-value<0.1, **...p-value<0.05, ***...p-value<0.01

ESTIMATING THE TPB IN A STRUCTURAL EQUATION MODEL

We test the Theory of Planned Behavior with a structural equation model (SEM) as in Carr and Sequeira (2007), Lortie and Castogiovanni (2015) or Zapkau et al. (2015), using ordinary least squares regressions. For the estimation, we first aggregate the relevant items of each factor. Applying (standardized) Cronbach's Alpha as a selection criterion, we identify which items should be used. The higher the value of Cronbach's Alpha, the more consistently the items capture the underlying theoretical construct. The measure should exceed 0.7, but as the calculation involves the number of items, the threshold declines in a setting with only few items per factor (Rosendahl Huber et al., 2014).

For the trainees of module 3 answering all questionnaires, Cronbach's Alpha ranges between 0.62 and 0.87 in QI, with the lowest values for the items of intention and highest for the attitude items. For the data collected in QIII, Cronbach's alpha increases strongly for the intention items and decreases for those of PBC, especially decision making under risk and probability of success. Overall, the consistency of most of the items increases after the course, such that we use all items for the calculation of the factors attitude, subjective norm and perceived behavioral control for the estimation of the structural model.

Figure 3 shows the share of participants with changing assessments between the surveys in QI and QIII for the aggregated items of attitude, subjective norm and PBC, and separately for the items of intention – the dependent variables of the subsequent regressions. Interestingly, the determinants of intentions show a larger variation between the beginning of the course and six months later than the measures of intention; especially the choice to become an entrepreneur is very stable. Taking into account the level of change as reported in Table 2 it appears that for the items of subjective norm and attitude, the positive and negative changes outweigh each other.

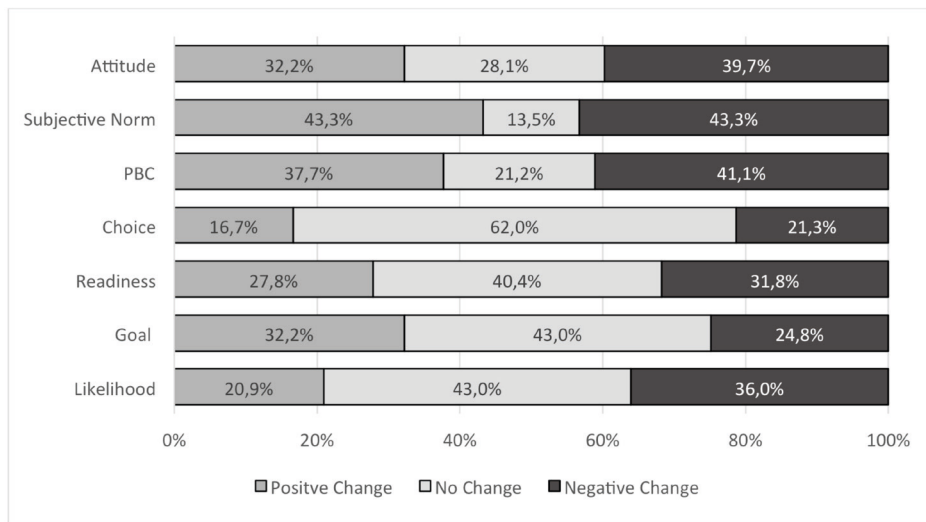


Figure 3: Percentage of respondents changing their assessment (QI-QIII, Module 3)
Source: Kägi et al. (2018), own calculations. N=158

The data allows us to perform an innovative approach in the TPB framework. To our knowledge, the analyses usually apply the TPB approach by using the absolute levels of the relevant items and factors. Our panel data structure allows us to use the relative change over time in the estimation. Thereby, we can observe the dynamics in the TPB, e.g. whether an increase in perceived behavioral control causes an increase in the intention to become an entrepreneur instead of looking at the determinants of entrepreneurial intentions in levels. In this sense, the model combines a before-after comparison with a structural equation model.

For the explanatory variables, we aggregate the change for each item and respondent between QI and QIII for the factors attitude, subjective norm, perceived behavioral control, and take the mean (average change per person). Further control variables (measured in levels) are the entrepreneurial background of the parents, work experience in a small or newly founded firm, age, gender and the region (serves as a university-dummy). Due to missing values due to the different survey applications across the Swiss regions, we are not able to properly control for the educational background.

Table 3 shows the regression results from OLS estimation with heteroscedasticity-robust standard errors for the intention items “my professional goal is to create a business someday” (panel 1) and “I am ready to do anything to be an entrepreneur” (panel 2). In panel 1, all three TBP-factors exhibit a significantly positive effect on the intention to create a business with the quantitatively largest effect for perceived behavioral control. Even when we include all three TPB-factors in the regression (column 4), the coefficient of PBC remains unchanged and is still significant. In columns 3 and 4 that control for the change in subjective norm, the dummy variable for females turns significant. This could indicate that once we take into account the perceived support from the close environment, females have less often the goal to become an entrepreneur (even among those who participate in the courses; their share is only about 25%). The effect of previous work experience is significantly negative in all regressions. An explanation is that a potential positive impact of previous work experience runs indirectly via the TPB-variables and that the remaining effect on intention is negative (Miralles et al, 2015).

Panel 1	My professional goal is to create a business someday			
	(1)	(2)	(3)	(4)
Female	-0.362	-0.464	-0.899**	-0.779**
Age	0.021	0.019	0.003	0.013
Entrepreneurial Parents	-0.032	0.053	0.143	0.089
Previous Work Experience	-0.457*	-0.490*	-0.737**	-0.651**
Perceived Behavioral Control	0.152***			0.151**
Attitude		0.074**		0.026
Subjective Norm			0.063*	0.033
N	133	134	92	86
Adjusted R2	0.114	0.071	0.141	0.251
F Statistic	3.432***	2.462**	3.137***	4.162***
Panel 2	I am ready to do anything to be an entrepreneur			
	(1)	(2)	(3)	(4)
Female	-0.091	-0.217	-0.366	-0.239
Age	0.011	0.008	-0.007	0.010
Entrepreneurial Parents	-0.071	-0.0001	-0.147	-0.186
Previous Work Experience	0.099	0.025	-0.192	-0.209
Perceived Behavioral Control	0.236***			0.200***
Attitude		0.135***		0.065
Subjective Norm			0.096***	0.040*
N	135	136	94	88
Adjusted R2	0.169	0.132	0.064	0.306
F Statistic	4.881***	3.922***	1.907*	5.256***

Table 3: Structural equation model of Module 3

Source: Kägi et al. (2018), own calculations

All TBP-variables are measured as the average sum of changes between QI and QIII; significant differences for p-values<0.1: *...p-value<0.1, **...p-value<0.05, ***...p-value<0.01 with robust standard errors. All regressions include regional dummies and a constant term.

In panel 2, only the TBP-factors have significant effects on changes in the intention to do anything to become an entrepreneur – the estimated coefficients are qualitatively and quantitatively similar to those in panel 1. Attitude and subjective norm exhibit, as in panel 1, as significantly positive effect when included separately in the regression but turn insignificant (less significant, panel 2) in column 4 due to multicollinearity. Overall, regarding the size and significance of the estimates as well as the share of explained variation, the estimation works best for the intention item readiness that shows the largest change between the two measurement points (the same holds for the intention-item likelihood, not reported in the table). Finally, robustness checks limiting the sample to those observations included in the regressions in column 4 show that the results are not driven by the sample size and composition. Using data collected in QII rather than QIII leads to smaller and largely insignificant estimates; this is plausible as the QII survey was conducted right after the course and the descriptive statistics show that the (full) change in the participants' assessments only appears with a time lag.

SUMMARY AND CONCLUSIONS

The paper presents the estimation results of a structural equation model using the Theory of Planned Behavior, performed in course of the impact analysis of the Innosuisse start-up training program (Kägi et al., 2018). The start-up training (called "CTI-Entrepreneurship Training Program" during the evaluation period) aims at supporting students and university members interested in setting up knowledge-intensive and technology-based companies by offering four different course modules. For modules two to four, the participants completed three surveys, the first at the beginning of the course, the second right after the course and the third six months after the course.

The panel structure of the data allows us to relate the change in the intention to become an entrepreneur to the changes in attitude, subjective norm and the perceived behavioral control of the course participants. The estimation results confirm the hypotheses of the Theory of Planned Behavior. The structural equation models show that an increase in the perceived behavioral control (i.e. the necessary skills) leads to a statistically significant increase in the intention to become an entrepreneur. This is particularly visible for Module 3, where participants already have tangible business models in their mind or are already in the process of setting up their business. Thus, they benefit from this five-days-course mainly by training their (perceived) skills. While also their attitude and the perceived support from the social environment change over time, these factors are at this stage less important for the formation of intentions.

Overall, if the course participation causes a change in the Perceived Behavioral Control, and these changes relate to the intention to become an entrepreneur, then we can conclude that the training program successfully contributes to the participants' being able to make a more informed choice about them becoming an entrepreneur, to enhance the entrepreneurial spirit for those who decide in favor of it, and to progress on their way toward fulfilling their intention by equipping them with the necessary skills. However, as the analysis does not involve a control group and thus is not able to account for unobserved heterogeneity between participants and non-participants, the results have to be interpreted with care.

REFERENCES

- Ajzen, I.** (1991). The Theory of Planned Behavior, *Organizational Behavior and Human Decision Processes* (50), pp. 107-122
- Carr, Jon C., and Sequeira, J.M.** (2007). Prior family business exposure as intergenerational influence and entrepreneurial intent. A Theory of Planned Behavior approach. *Journal of Business Research* 60, pp. 1090-1098
- Heuer, A., Kolvereid, L.** (2013). Education in entrepreneurship and the Theory of Planned Behaviour, *European Journal of Training and Education*, 38(6), pp. 506-523
- Kägi, W., Oswald, A., Suri, M., Kaufmann, P., and Nindl, E.** (2018). CTI Entrepreneurship - Wirkungsanalyse und Evaluation. Studie durch BSS Volkswirtschaftliche Beratung und KMU Forschung Austria im Auftrag von Innosuisse, Bern
- Liñán, F., Fayolle, A.** (2015). A systematic literature review on entrepreneurial intentions: citation, thematic analyses, and research agenda, *International Entrepreneurship and Management Journal*, pp.907-933
- Lortie, J., and Castogiovanni, G.** (2015). The theory of planned behaviour in entrepreneurship research: what we know and future directions. *International Entrepreneurship and Management Journal* 11, pp. 935-957
- Miralles, F., Giones, F., and Riverole, C.** (2016). Evaluating the impact of prior experience in entrepreneurial intention. *International Entrepreneurship and Management Journal* 12(3), pp. 791-813
- Osterbeek, H., van Praag, M., and Ijsselstein, A.** (2010). The impact of entrepreneurship education on entrepreneurship skills and Motivation. *European Economic Review* 54, pp. 442-454
- Rosendahl Huber L., Sloof, R., and van Praag, M.** (2014). The effect of early entrepreneurship education: Evidence from a field experiment. *European Economic Review* 72, pp. 76-97
- Zapkau, F.B., Schwens, C., Steinmetz, H., and Kabst, R.** (2015). Disentangling the effect of prior entrepreneurial exposure on entrepreneurial intention. *Journal of Business Research* 68, pp. 639-653

AUTHORS

ELISABETH NINDL

E: e.nindl@kmuforschung.ac.at

PETER KAUFMANN

E: p.kaufmann@kmuforschung.ac.at

KMU Forschung Austria – Austrian Institute for SME Research
Gusshausstraße 8, 1040 Vienna

KEYWORDS:

Entrepreneurship training, Theory of Planned Behavior, structural equation model

USING THE DELPHI METHOD IN EVALUATIONS – INCORPORATING A FUTURE ORIENTED PERSPECTIVE IN EVALUATIONS

MICHAEL DINGES, ANNA WANG AND KLAUS SCHUCH

DOI: 10.22163/fteval.2020.469

ABSTRACT

This article discusses the scope of use of the Delphi method in evaluations. It suggests that the Delphi method is a promising approach for validating the evidence gathered by incorporating a forward-looking perspective in the drawing of conclusions and in elaborating recommendations by making use of expert knowledge from different domains. Thereby, the Delphi method also facilitates information and engagement of relevant stakeholder communities in an ongoing evaluation process. Based upon the application of the method in the Evaluation of the Support Structures for Horizon 2020, EUREKA, COSME, EEN and ERA in Austria, the article demonstrates how the application of the Delphi method allows to analyse the degree of fit of an existing service portfolio against changing framework conditions and expected developments in the future.

INTRODUCTION

Evaluations of R&I interventions face the challenge of transforming evidence on the performance of an intervention into recommendations on design options and broader policy recommendations in the observed domain. In the process of formulating recommendations, the core perspective of an evaluation changes: While the analysis of performance is predominantly backward oriented by considering the constituent actors and their changing contexts in an ex-post perspective, the process of formulating recommendations and design options requires careful consideration of the existing evidence combined with possible future developments of contexts and actors addressed by an intervention. Thus the perspective changes from retrospective (evidence-based answer to questions on ‘what happened’ and ‘what worked well or not’) to prospective (‘what likely needs to be done to anticipate and successfully tackle future challenges’).

While evaluations are usually categorized as being largely formative or summative (Chen 1996, Patton 1996), Edler et al. (2012) show in their empirical analysis that: 1) rather summative issues such as ‘outputs, outcomes, and impacts’, ‘goal attainment and effectiveness’, ‘internal and external consistency’ in evaluations are strongly linked to more formative topics such as ‘policy/strategy development’ and 2) this type of “holistic evaluations” seeks to understand and measure the programme

logic. This type of evaluation usually focuses on combining survey data with case studies, some (limited) network analysis and applies, as a consequence of the additionality assessment, some (limited) before/after group comparisons (ibid.).

However, even holistic evaluations performed today fall short in applying methods, which focus on (1) the possible future evolution of relevant contexts of an intervention as well as (2) presumed strategies of different types of actors towards the anticipated futures. Still, evaluation methods and monitoring practices are largely based on linear ideas and most evaluations are backward looking (ex post) despite the development and application of ex ante evaluation methods (Nieminen and Hyytinen, 2015). In particular, the application of methods, which can provide future oriented strategic decision support, are yet untapped in current evaluation practices of innovation policies.

Against this background we ask how a forward-looking perspective can be integrated in evaluations in a cost-efficient and effective manner? Nieminen and Hyytinen (2015) suggest that the essential characteristics of evaluation and foresight methodologies, complemented by dynamic system modelling as well as robust societal embedding, can provide both high-level ‘visionary’ inputs together with specific and detailed information for decision-making.

Taking up this perspective, we demonstrate how the Delphi method can be used as a tool not only for validating the evidence gathered, but also how to incorporate a forward-looking perspective in drawing conclusions and elaborating recommendations by making use of expert knowledge from different domains. The empirical basis for the discussion is the Evaluation of the Support Structures for implementing the European Framework Programme for Research and Innovation (H2020), EUREKA, COSME, EEN and ERA in Austria (Dinges et al. 2018) commissioned by the Austrian Federal Ministry of Science, Research and Economy in 2017.

We first provide a short conceptual outline on the scope and use of the Delphi Method as referenced in the academic literature. We then proceed to describe how the Delphi-Process was implemented in the Evaluation of the Support Structures for Horizon 2020 in Austria. We show how the findings of the Delphi process influenced the process of synthesizing evaluation results and how policy recommendations were shaped through the methodological design. In a final section we discuss how the Delphi process can make evaluations more sensitive towards changing framework conditions and uncertain futures, and which prerequisites for implementation have to be taken into account.

THE DELPHI METHOD

The Delphi Method was originally developed by the RAND Corporation for technological forecasting and applying a systematic approach to the utilization of expert opinions. The Delphi Method attempts to make use of informed intuitive judgement (of experts) and derives its importance from the realization that projections into the future, on which policy decisions must rely, are largely based on the personal expectations of individuals rather than on predictions derived from a well-established theory (Helmer, 1967)¹.

In its original configuration, the Delphi Method sets out a process, in which the various opinions of several experts on a particular issue, are combined into a single combined position. It is a method for structuring a group communication process so that the process is effective in allowing a group of individuals, as a whole, to deal with a complex problem (Linstone and Turoff, 1975; Cuhls, Blind and Grupp, 2002). Key objectives of the Delphi Method were to reduce the negative effects of group interactions (Gupta and Clarke, 1996) and to obtain the most reliable consensus of opinion of a group of experts (Dalkey and Helmer, 1962). Key design elements of the Delphi Method are anonymity, iteration, controlled feedback and participating experts (Geist, 2009; Rowe and Wright 1999):

- The anonymity of the Delphi allows all panellists to be removed from pressures encountered in a face-to-face interaction. All ratings and comments are submitted anonymously, therefore, members can change their minds without feeling judged by others in the group (Rowe et al., 1991).
- Group members, or panellists, generate ideas and comments about the issue or problem from individual brainstorming. The researcher distils those responses and presents them to the panellists in the form of a survey for a second round of input. Then, the results and responses are presented to the group members in condensed form for comment. This process is being iterated up to four times.
- Controlled feedback occurs between iterations when the researcher uses qualitative data (e.g. comments, reasons for ratings, etc.) as a form of qualitative feedback. Controlled feedback presented in an organized format allows panellists to read, comment on, and critique all the facets of the issue virtually simultaneously between iterations (Rowe et al., 1991).
- Statistical group response consists of quantitative feedback (e.g. medians and interquartile ranges, or means and standard deviations) based on the numerical ratings of each item. After the final iteration, the ideas and opinions are listed along with the descriptive statistics of the ratings (Rowe et al., 1991).

Over the decades, a number of variations of the method have been developed, adapting the application in different ways, such as techniques to select participants, types of questions employed, tools used for the analysis of responses, and type of outcome sought (Hirschhorn 2019). The application possibilities for Delphi processes were also widened through advances in ICT support tools, allowing for example to implement real-time Delphis and experimentation by securing the anonymity of participants etc. Brady (2015) details that the Delphi Method has been used in an array of different contexts, where expert knowledge is needed to inform decision making or to understand a phenomenon in

greater depth including in particular public policy, health care, education, and technological forecasts. In public policy, the Delphi Method seeks to provide policy makers with a better understanding of policy design and implementation (Alder and Ziglio, 1996; Linstone and Turoff, 1975).

THE USE OF DELPHIS IN R&I POLICY AND R&I POLICY EVALUATION

In the domain of R&I policy, the focus of the use of Delphis changed in the last decades. Whereas in the 1970s and 80s, Delphi exercises and critical technology studies conducted in the US, in France and the Netherlands focused on predicting and controlling the future, new forms of Delphis have been developed that do not necessarily strive to achieve consensus on future forecasts, but rather to map the diversity of opinion in order to enable public policy to imagine alternative paths of future developments and to design public policy in a responsive manner towards these potential paths (Havas, Schartinger and Weber 2009). For example, the foresight study in support of future EU Research and Innovation Policy (BOHEMIA, Weber et al. 2018), aimed to support the deliberations on future proposals for a post-2020 Research and Innovation Framework. The BOHEMIA Delphi statements were not meant to provide a broad overview of 'all things to come' but started from scenarios on the Future for Research and Innovation Policies in Europe.

For monitoring and evaluation, the Delphi technique can be used in various evaluation tasks to anticipate possible futures, for instance by assessing the 'problem' a certain program aims to solve and forecasting future changes². It is a particular useful method in judgment and forecasting of situations in which pure model-based statistical options are not practical (ibidem). Despite its potential usefulness, the Delphi Method has not been widely used in evaluations of R&I policies and institutions.

In the following, we therefore make the case how the Delphi technique was used in the case of the evaluation of the Austrian support structures for Horizon 2020, EUREKA, COSME, EEN and ERA.

THE APPLICATION OF THE DELPHI METHOD IN THE EVALUATION OF THE SUPPORT STRUCTURES FOR HORIZON 2020 IN AUSTRIA

SCOPE AND DESIGN OF THE EVALUATION

The evaluation aimed to undertake a systematic assessment of the Austrian support structures and implementation of Horizon 2020 and ERA, as well as EUREKA, COSME, and the Enterprise Europe Network

1 <https://www.rand.org/content/dam/rand/pubs/papers/2008/P3558.pdf>

2 <https://www.betterevaluation.org/en/evaluation-options/delphitechnique>

(Dinges et al. 2018). The evaluation focused on the assessment of support services provided by the Austrian Research Promotion Agency – Division of European and International Programmes (FFG-EIP), which is the national contact point for Horizon 2020 on behalf of the Austrian Federal Ministry of Education, Science and Research³. FFG-EIP pursues the following objectives:

- to support successful participation in Horizon 2020 by means of information and advice,
- to provide system-orientated support for Austrian R&I organisations,
- to be an expert partner for European R&I programmes and ERA developments, and
- to provide analysis on participation and to provide R&I policy support.

Against these objectives of FFG-EIP, the evaluation commissioned by the ministry focused on the following areas:

- General evaluation of the implementation of the support structures for Horizon 2020: Are orientation and implementation of available formats and support services - within the allocated budget - suitable and appropriately implemented to achieve the desired effects of measures regarding output as efficiently and effectively as possible?
- Evaluation of specific services: By respecting observable and anticipated developments in the R&I domain (including R&I policy), are the service formats adequate and which improvements should be made?
- Action required and options for action: Which white spots can be encountered and which new and better approaches and tools should be established? According to which principles should these essentially be designed?

While having a predominantly summative and ex-post oriented character, the evaluation questions emphasized that also anticipated developments in the R&I domain should be taken into account and that the adequacy of support formats needs to be assessed against current and future developments in the European and Austrian research and innovations system. Future developments, for example related to the European R&I policy domain and in particular to the development of the new Framework Programme for R&I, Horizon Europe, as well as the future shaping of the European Research Area, but also to the development of strategic capacities of R&I actors, should be considered.

The evaluation approach was set up against an impact model, which linked actions of FFG-EIP with desired outcomes and impacts. For answering the evaluation questions, a mixed method evaluation design was chosen. It allowed for triangulation of quantitative and qualitative methods. Key components included online surveys with key stakeholder groups, interviews with a broad range of R&I policy actors and stakeholders, focus group interviews, the Delphi Method, validation workshops, and critical commenting through an external advisory board.

Within this overall design, the two-stage online Delphi Method was chosen as a key component in the validation of evaluation results and the development of scenarios and policy recommendations by providing converging expert assessments on hypotheses and preliminary conclusions during the final phase of the evaluation.

The evaluation was performed from August 2017 until end of June 2018. Within this timeframe, the two rounds of the Delphi survey were operationalised from February until beginning of May 2018.

SELECTION OF EXPERTS

In order to benefit from a broad range of expert assessments on the support services and governance for Horizon 2020 in Austria, a sample of 190 experts were chosen due to their high knowledge of the evaluation subject at hand. They are formed from the following stakeholder groups: 1) senior university leadership (rectors and vice-rectors), 2) management of non-university research organizations, 3) management of enterprises that are highly active in Horizon 2020, 4) policymakers, including delegates and experts of Austria to Horizon 2020 programme committees and representatives of relevant ministries, 5) representatives of national and regional funding agencies.

To ensure an adequate range of opinions, the experts were approximately equally distributed across the five stakeholder groups. In the first stage, 69 experts shared their assessment and were subsequently invited to the second-stage. 39 of these responded to the second round of the Delphi.

IMPLEMENTATION OF THE DELPHI

The Delphi surveys were part of the final phases of the evaluation, when other forms of data collection and analysis have had already been concluded. Thus, it was applied as a method of gathering expert assessments on specific hypotheses and conclusions derived from the evaluation in order to feed into a forward-looking development of scenarios and recommendations. Five overarching thematic areas were covered, each with their corresponding hypotheses and conclusions. These topics include:

1. the quality, speed, and transparency of communication and information on developments in European R&I,
2. the assumed necessity for strategically supporting companies, universities and research organizations in developing (pro-)active participation strategies for Horizon 2020,
3. the increasing demand for improved strategic support services in light of new thematic and organisational challenges posed by Horizon Europe,
4. the assumed importance of increased strategic multi-level coordination of stakeholders for European public-public and public-private partnership instruments,
5. the expected impacts of the new mission-oriented approach in Horizon Europe for national R&I policies and governance.

Each topic was introduced by describing its context and importance. In addition, likely demands placed on R&I governance and support services – based on expected developments in Austrian and European R&I policy – were shortly mentioned. In the first stage of the Delphi, a set of three to five hypotheses were assigned to each topic. These hypotheses were based on the findings of the previous analytical evaluation steps

and formulated as preliminary conclusions or recommendations to R&I policymakers. The participating experts were asked to rate each quantitatively using a five-point Likert scale from “Yes, I fully agree” to “No, I strongly disagree”. Additionally, experts were asked to give a qualitative

statement explaining their rating for each. Table 1 Example of Delphi stage 1 questions provides an example of how a topic was presented, as well as how both open-ended and closed questions were utilized in the Delphi.

Topic: Mission-orientation and preparation for Horizon Europe
<p>Hypothesis:</p> <p>Horizon Europe will most likely be characterized by a stronger mission-orientation of R&I policy. Thus, the new framework programme will have a set of pre-defined cross-thematic and cross-disciplinary R&I goals for funding. These goals should have high societal impacts (e.g., at the interface of energy, environment, mobility, health, security, etc.) that should be measurably achieved in a relatively short time frame and carry with them significant value for society. This new mission-oriented approach to R&I policy demands: i) a common understanding of specific joint goals, ii) a strong alignment of European and national R&I policies, iii) strong governance and strategic mission management on programme and project levels. Therefore, mission-orientation imposes high demands on the formulation, implementation, and support of R&I policy.</p>
<p>Please assess the following statements:</p>
<p>Statement 1: Innovation is not sufficiently important in the European Research Area.</p>
<p> <input type="radio"/> 1 – No, I strongly disagree <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5 – Yes, I fully agree </p>
<p>Please explain your answer:</p>
<p>Statement 2: In case there is a stronger mission-orientation on European level, Austria needs a national mission strategy, i.e., a with national stakeholders coordinated common approach.</p>
<p> <input type="radio"/> 1 – No, I strongly disagree <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5 – Yes, I fully agree </p>
<p>Please explain your answer:</p>
<p>Statement 3: With a stronger mission-orientation on European level, Austrian R&I funding needs to be adjusted to align with European missions.</p>
<p> <input type="radio"/> 1 – No, I strongly disagree <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5 – Yes, I fully agree </p>
<p>Please explain your answer:</p>
<p>Statement 4: With a stronger mission-orientation on European level, the Austrian Research Promotion Agency needs to improve its support services in information provision, community management, and coordination.</p>
<p> <input type="radio"/> 1 – No, I strongly disagree <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5 – Yes, I fully agree </p>
<p>Please explain your answer:</p>

Table 1 Example of Delphi stage 1 questions

ANALYSIS OF RESULTS OF THE FIRST STAGE

Overall, experts agreed moderately to strongly with the hypotheses of the first stage. Additionally, the use of free text answers for each statement under scrutiny allowed them to elaborate on their quantitative assessment. This feedback provided valuable input and explanations for (dis-)agreement and highlighted ideas for areas of improvement in R&I policy design. Thus, it can be argued that the first stage of the Delphi survey can be characterized as serving two key functions: i) validation of evaluation findings and preliminary conclusions, and ii) scoping forward-looking options for R&I policy design in response to expected changes at European and national R&I policy level.

The results of the first stage were compiled, analysed, synthesised and presented to the expert group during the second Delphi stage: First, the quantitative ratings were calculated for each question and summarized in a bar chart. Then the free text answers were analysed and summarized and, if relevant differences became apparent, differentiated by stakeholder groupings. The aim was to synthesize all answers into a 'story' to present the range of opinions, if and why the hypothesis was (in-)correct, and whether any important aspects were missing. The results of both the quantitative and qualitative analysis were then used to re-formulate hypotheses into conclusions by taking into account inputs and ideas raised by the experts.

THE SECOND STAGE OF THE DELPHI

In the second round of the Delphi survey, the respondents of the first round were confronted with the summative results of the first round in the forms of a graphical representation of quantitative assessments, the summaries of open-ended answers, as well as re-formulated conclusions, sometimes complemented by recommendations.

The formulation of new questions closely resembled those of the first stage by making use of both quantitative and qualitative assessment options: For each conclusion, participants were asked to give again a quantitative assessment on a five-point Likert scale and were invited to take the opportunity to comment.

SYNTHESIS OF RESULTS

The results of the second stage showed that expert opinions converged to a certain degree. Most of the responses displayed high agreement with our conclusions. The analysis and synthesis of results followed the same two-step method of summarizing quantitative results, drafting a summarized overview of qualitative responses, and, if necessary, re-formulating conclusions while incorporating additional points raised by surveyed experts.

As a result of the two-stage process, the Delphi produced a set of validated conclusions that took into account anticipated demands on changes in policy orientation and governance associated with future developments at the level of national and European R&I policy. The Delphi survey showed in particular, that the future role, functions and tasks of FFG-EIP do not only depend on the framework programme - they also depend on Austria's intended positioning vis-a-vis European R&I policy.

POLICY PATHWAYS FOR AUSTRIA IN VIEW OF HORIZON EUROPE

Based on the results of the Delphi survey three alternative policy pathways for the future orientation of Austrian governance and support services for Horizon Europe were synthesized and finally discussed with R&I policy-makers in Austria. These pathways were in particular developed in order not to limit the further design of the roles, functions and tasks of the FFG-EIP to rectify existing deficiencies or just to improve them against the backdrop of current tasks and challenges, but to take into account future challenges stemming from changing framework conditions such as the EC's proposal for the 9th European Framework Programme for Research and Innovation ("Horizon Europe") (see Dinges et al, 2018 for further details).

Our evaluative findings about the expectations of the Austrian R&I community clearly showed that changes will be necessary in governance at national level and also at the level of the interface with European policy, as well as in the support services for the R&I community. It became also clear that the extent of these changes depends on Austria's intended positioning regarding European R&I policy. The three pathways reflect options for action at national level and the respective consequences that these would have for the support structures. They were framed as "food for thought" for further discussion in Austria, not least, also with regard to the development of a new R&I strategy.

The three following scenarios/policy pathways were constructed:

1. "Enhancement in continuity" scenario: Incremental change
2. "Smart and proactive alignment" scenario: R&I policy as a catalyst in the European multilevel system
3. "Distributed empowerment" scenario: Strengthening and cross-linking the R&I community

The central idea of the first scenario is that the well-established and largely well-functioning status quo is maintained and incrementally adapted to the requirements stipulated by the new tools and initiatives at European level, but without significant changes having to be made in Austria, e.g. in the area of governance. Consequently, elements of national strategy planning and prioritisation (e.g. with regard to Austrian participation in strategic partnerships) is negligible or limited to individual preferences of ministries.

In the second scenario, Austrian policy develops its role to that of a (pro)active contributor to European R&I policy, and to make more targeted and effective use of the opportunities resulting from this for Austria ('smart alignment'). To be better able to positioning Austria as an integral element of the multilevel innovation ecosystem in Europe, Austria would need to implement a range of changes within the national system to achieve more intensive and more influential cooperation in the design and subsequently in the use of European R&I policy. Specifying such a concept for a consistent Austrian positioning within the overall European R&I policy would need to be a central element within the forthcoming Austrian R&I strategy.

The third scenario sees the extension of participation and involvement in European R&I policy, particularly by means of empowering the strategy and networking capabilities of R&I actors themselves. This scenario assumes that organisations have an independent development strategy and ability to self-organisation in strong networks. However, for this policy-pathway also appropriate resources and capacities are requi-

red, whose mobilisation needs to be supported by R&I policy (e.g. by means of financial incentives, networking activities, information sharing, etc.). The experiences of other European countries show that there can be losers as well as winners in this empowerment process.

DISCUSSION AND CONCLUSION: THE UTILITY OF THE DELPHI METHOD IN R&I POLICY EVALUATION

As demonstrated by the case of the Evaluation of the Support Structures for Horizon 2020, EUREKA, COSME, EEN and ERA in Austria, the Delphi Method can be a useful tool for supporting the formulation of policy recommendations and pathways that consider challenges and potentials for re-orientation of R&I policies that are associated with future developments.

The case reveals that the implementation of a Delphi in evaluation of R&I policy must be well-timed in the overall research design, and best applied in the final phases of an evaluation with the aim to produce a set of validated, future-oriented conclusions and recommendations.

The combination of both a quantitative and a qualitative element in each question allows the Delphi Method to be particularly useful for the validation of evaluation findings while at the same time serving to gather expert ideas on improvements in R&I policy design.

Special care needs to be taken in the selection of experts to ensure the reflection of a broad range of opinion and expertise, as this might have a high impact on the quality of results. Due to the efforts required for the analysis of qualitative data, the sample should be limited in size but include all relevant experts and stakeholder groups in an appropriate distribution. The relatively high statistical mortality in running at least two rounds of Delphi remains a problematic issue and needs to be considered respectively. In our case the statistical mortality was in total almost 80%. This resulted in an above-average "survival" of stakeholder groups that are more used to formulate and to take positions (i.e. policy-makers and representatives of large companies and interest groups), which can lead to a distortion of perceptions.

To reduce statistical mortality, the effort for the addressed experts should be kept as limited as possible without destroying the overall narrative of the formulated hypothesis and the questions derived thereof. Otherwise the Delphi exercise would run risk to transform into a series of hardly connected and de-contextualized surveys.

Thus, it maybe would have been better in our specific case to confront the addressed experts from the outset with a couple of scenarios/policy pathways to which the hypothesis should have explicitly referred to. Such a scenario-based approach, which is broken down to the level of specific questions, would probably have provided more clear narratives and orientation and kept the attention of the experts for a longer duration, thus reducing statistical mortality. The approach taken by us, however, was in the reverse direction: the alternative future policy pathways for the Austrian R&I policy to approach the challenges of the next Framework Programme for R&I were constructed after the Delphi and not used as input to the Delphi although ideas of such policy pathways were implicitly already available, but neither transparently formulated nor explicitly referred to.

Despite the challenge of statistical mortality, the Delphi Method is particularly useful in situations where a broad range of stakeholders might be affected differently or place different demands on R&I policy design in response to expected changes. It is a structured means of interaction in an evaluation process, which also allows to present and discuss core findings of an evaluation with a broader community of stakeholders.

In a more general perspective, the Delphi Method can be seen as a stakeholder engaging method, which allows to 1) put stronger emphasis on changing framework conditions in the synthesis of evaluation results, 2) facilitates to incorporate ex-post findings with a more forward looking perspective, and 3) provides better legitimization of recommendations.

Compared with large, iterative stakeholder workshops, the Delphi Method in R&I policy evaluation facilitates the formulation of targeted, forward-looking policy options and recommendations that synthesize the collective expertise and knowledge available in a more time- and cost-efficient manner. Also feedback from more remote respondents (in terms of geography) can be included, who would usually rather not come to a workshop organised in the capital city Vienna. The results of the Delphi survey can be further fed into the discussion process with the contractor and a number of core stakeholders where they can contribute to increase acceptance of evaluation results and increase understanding for possible policy options.

REFERENCES

- Adler, M., and Ziglio, E.** (Eds.). (1996). *Gazing into the oracle*. Bristol, PA: Kingsley.
- Brady, S. R.** (2015). Utilizing and Adapting the Delphi Method for Use in Qualitative Research. <https://doi.org/10.1177/1609406915621381>
- Chen, H-T** (1996). A comprehensive typology of program evaluation. *Evaluation Practice*, Volume: 17 issue: 2, page(s): 121-130. Issue published: June 1, 1996; <https://doi.org/10.1177/109821409601700204>
- Cuhls, K., Blind, K. and Grupp, H.** (2002). *Innovations for our Future. Delphi '98: New Foresight on Science and Technology*. Series: Technology, Innovation and Policy (ISI). Physica-Verlag.
- Dalkey, N. and Helmer, O.** (1962). An experimental application of the Delphi Method to the use of experts. Prepared for the US Air Force Project Rand. Manuscript. RAND Cooperation.
- Dinges, M. et al.** (2018). Evaluierung der Umsetzung von H2020, EUREKA, COSME, EEN und ERA in Österreich. 10.22163/fteval.2018.291.
- Edler, J., Berger, M., Dinges, M. and Gök, A.** (2012). The practice of evaluation in innovation policy in Europe. *Research Evaluation*, 21 (3), 167-182.
- Geist, M.** (2009). Using the Delphi Method to engage stakeholders: A comparison of two studies. *Evaluation and program planning* 33(2):147-54. DOI: 10.1016/j.evalprogplan.2009.06.006
- Gupta, U. G., and Clarke, R. E.** (1967). Theory and applications of the Delphi technique: A bibliography (1975–1994). *Technological Forecasting and Social Change*, Volume 53, Issue 2, October 1996, Pages 185-211.

Havas, A., Schartinger, D. and Weber, M. (2010). The Impact of Foresight on Innovation Policy-Making: Recent Experiences and Future Perspectives. *Research Evaluation*, 19(2):91-104. DOI: 10.3152/095820210X510133

Helmer, O. (1967). *Analysis of the Future. The Delphi Method.* Manuscript. The RAND Corporation.

Hirschhorn, F. (2019). Reflections on the application of the Delphi Method: lessons from a case in public transport research. *International Journal of Social Research Methodology*, 22:3, 309-322, DOI: 10.1080/13645579.2018.1543841.

Linstone, H.A. and Turoff, M. (1975). *The Delphi Method: Techniques and Applications*, Addison-Wesley Publ., Reading/Mass.

Nieminen, M., and Hyttinen, K. (2015). Future-oriented impact assessment: Supporting strategic decision-making in complex socio-technical environments. *Evaluation*, 21(4), 448-461.

Patton, M. Q. (1996). A world larger than formative and summative. *Evaluation Practice*, Volume: 17 issue: 2, page(s): 131-144. Issue published: June 1, 1996; <https://doi.org/10.1177/109821409601700205>

Rowe, G. and Wright, G. (1999). The Delphi technique as a forecasting tool: issues and analysis. *International Journal of Forecasting*, 15 (1999) 353–375.

Rowe, G., Wright, G. and Bolger, F. (1991). The Delphi technique: a re-evaluation of research and theory. *Technological Forecasting and Social Change*, 39(3), 235-251.

Weber et al. (2018). *Transitions on the Horizon: Perspectives for the European Union's future research and innovation policies.* Final report from project BOHEMIA - Beyond the horizon: Foresight in support of the EU's future research and innovation policy. Luxembourg: Publications Office of the European Union. doi: 10.2777/493572.

AUTHORS

MICHAEL DINGES

Austrian Institute of Technology
Giefinggasse 4, 1210 Wien
E: michael.dinges@ait.ac.at

ANNA WANG

Austrian Institute of Technology
Giefinggasse 4, 1210 Wien
E: anna.wang@ait.ac.at

KLAUS SCHUCH

Centre for Social Innovation
Linke Wienzeile 246, 1150 Vienna
E: schuch@zsi.at

WIE KÖNNEN WIDERSPRÜCHE ZWISCHEN MARKTDURCHDRINGUNG UND REBOUND-VERMEIDUNG GELÖST WERDEN?

SYSTEMANALYSE KONVERGENTER UND DIVERGENTER EINFLUSSFAKTOREN AN DEN BEISPIELEN E-AUTO UND GEBÄUDESANIERUNG

VERONIKA KULMER, SEBASTIAN SEEBAUER UND CLAUDIA FRUHMANN

DOI: 10.22163/fteval.2020.470

ZUSAMMENFASSUNG

Politikstrategien zur kohlenstoffarmen Transformation treiben die breite Markteinführung energieeffizienter Technologien voran, übersehen aber oft die damit verbundenen Reboundeffekte. Der Beitrag illustriert grundlegende Konflikte zwischen Technologieadoption und –nutzung an den Beispielen E-Autos und Gebäudesanierung. Systemmodelle führen die Ergebnisse von drei komplementären Forschungsmethoden zusammen: fuzzy cognitive mapping von Expertenwissen, Haushaltsbefragung und makroökonomische Analyse. Indem Politikimpulse durch das Wirkungsgeflecht verschiedener Einflussfaktoren verfolgt werden, können Hebelelemente für gezielte politische Maßnahmen bestimmt werden. Konvergente Einflussfaktoren (z.B. Umweltwerte) begünstigen den Markteintritt der Technologie und schwächen gleichzeitig Reboundeffekte ab. Divergente Einflussfaktoren (z.B. Einkommen) beschleunigen die Technologieadoption, ermöglichen aber auch späteren Rebound. Die Vermeidung von indirektem Rebound ist kritisch für die Realisierung von CO₂-Einsparungen. Anstelle von Einzelmaßnahmen sind Maßnahmenbündel zu empfehlen, die mehrere Wirkungskanäle ansprechen, etwa durch eine Kombination von fiskalpolitischen Instrumenten, bewusstseinsbildenden Kampagnen und Schulungen von Marktakteuren.

ABSTRACT

Low carbon transformation policies advocate the market diffusion of energy efficient technologies, but tend to overlook associated rebound effects. For the case of e-cars and building insulation in Austria we address inherent conflicts between technology adoption and usage. System models integrate results of three complementary research methods: fuzzy cognitive mapping of expert knowledge, household survey and macroeconomic analysis. By tracing policy impulses through the web of interlinked drivers, the system models illustrate leverage for targeted

interventions. Convergent drivers (e.g. pro-environmental values) stimulate the market uptake of the technology, and simultaneously constrain rebound effects. Divergent drivers (e.g. income) promote adoption, but set the stage for ensuing rebound effects. Preventing indirect rebound is found most critical to achieve energy savings. Instead of stand-alone measures, policy mixes should address multiple impact paths, e.g. by combining fiscal instruments, awareness building and training of key market actors.

EINLEITUNG

Um die ambitionierten Klimaziele zu erreichen, die 2015 bei der COP21 in Paris beschlossen wurden, setzen zahlreiche Länder auf innovative, energieeffiziente Technologien. Die breite Markteinführung von E-Autos oder die Sanierung des Gebäudebestands durch Wärmedämmung sollen Energienachfrage und Treibhausgasemissionen senken. Diese politische Strategie ist erfolgreich, wenn einerseits energieeffiziente Technologien schnell den Massenmarkt durchdringen, und andererseits technische Effizienzpotenziale nicht durch veränderte Nutzung kompensiert werden. Der vorliegende Artikel argumentiert, dass diese beiden Prozesse, Adoption und Rebound-Effekt durch veränderte Nutzung, von denselben Faktoren beeinflusst werden. Klimapolitische Instrumente, mit dem Ziel die Adoptionsrate zu steigern, legen oft den Grundstein für einen späteren Rebound-Effekt. Beispielsweise wird der Kauf eines E-Autos durch die Aussicht auf reservierte E-Parkplätze in Innenstädten attraktiver (Holtsmart & Skonhott 2014). Die gute Parkplatzverfügbarkeit legt es dann aber den NutzerInnen nahe, das E-Auto auch auf Wegen zu verwenden, die früher mit dem Fahrrad oder mit öffentlichen Verkehrsmitteln zurückgelegt wurden. Gebäudesanierung wird mit der Aussicht auf geringere Heizkosten beworben. Je mehr diese Motivation im Vordergrund steht, desto stärker wird den BewohnerInnen nahegelegt, sich mit dem billigeren Heizsystem eine komfortablere Raumtemperatur zu gönnen (de la Rue Can et al. 2015, Font Vivanco et al. 2016).

Adoption und Nutzung sind durch mehrere Facetten miteinander verknüpft. Zum Zeitpunkt der Kaufentscheidung für eine energieeffiziente Technologie antizipiert man bereits, wie die Technologie voraussichtlich genutzt wird – wie gut man damit seine alltäglichen Bedürfnisse befriedigen kann, wann sich die Investitionskosten amortisieren, wie hoch die laufenden Kosten sind, etc. (Wolf & Seebauer 2014). Da die effiziente Technologie die Energiedienstleistung mit weniger Energieeinsatz und folglich geringeren Kosten bereitstellt, steigt einerseits die Nutzung (direkter Rebound-Effekt; Sorrell 2007, Santarius 2014). Andererseits steigt durch die Kosteneinsparung das verfügbare Einkommen, welches für andere energieverbrauchende Güter und Dienstleistungen ausgegeben werden kann (indirekter Rebound-Effekt; Azevedo 2014, Thomas und Azevedo 2013). Diese Anpassungsprozesse von Angebot und Nachfrage summieren sich über alle Wirtschaftssektoren (gesamtwirtschaftlicher Rebound-Effekt; Allen et al. 2007, Turner 2013, Gillingham et al. 2016). Rebound-Effekte auf NutzerInnenebene entstehen nicht nur durch Preis-anreize, sondern auch durch Moral Licensing: Mit der Adoption der effizienten Technologie hat man das Gefühl, bereits einen ökologischen Beitrag geleistet zu haben. Nun darf man sich ohne schlechtes Gewissen einen Mehrverbrauch in anderen Konsumbereichen gönnen (Friedrichsmeier & Matthies 2015).

Die Erkenntnis, dass erwartete Emissionseinsparungen durch energieeffiziente Technologien oft nicht vollständig realisiert werden, wird vermehrt in der politischen Diskussion aufgegriffen (Font Vivanco et al. 2016). Direkte Rebound-Effekte in Wohnen und Verkehr werden auf 10-30% der erwarteten Einsparungen geschätzt (Sorrell 2007). Die Bandbreite der geschätzten indirekten und gesamtwirtschaftlichen Rebound-Effekte liegt zwischen 20% und 300% (Allan et al. 2007, Guerra & Sancho 2010, Turner 2009). Angesichts dieser Höhe von Rebound-Effekten besteht dringender Bedarf, politische Handlungsoptionen zur Rebound-Prävention zu entwickeln.

Sowohl Adoption als auch (veränderte) Nutzung finden im selben soziotechnischen Regime, in Interaktion mit den gleichen Marktakteuren und beeinflusst von denselben Personenmerkmalen statt (Geels 2004). Anschaffungs- und Betriebskosten, vorhandene Infrastruktur und konkurrierende Produkte am Markt sind sowohl für die Kaufentscheidung als auch für die laufende Nutzung relevant (Boulanger et al. 2013). Subventionen und Verordnungen sowie Kommunikation in den Massenmedien spielen ebenfalls eine doppelte Rolle (Steg et al. 2015). Sowohl die Kaufentscheidung als auch die alltägliche Nutzung sind abhängig von Einkommen, Umweltwerten oder Wissensstand der KonsumentInnen (Peters et al. 2012). Ein Politikinstrument das beispielsweise bei der Infrastruktur ansetzt oder Anreize für einzelne Marktakteure wie z.B. EinzelhändlerInnen und HandwerkerInnen bereitstellt, wird daher sowohl Adoption als auch Rebound beeinflussen.

Bis jetzt wurden Einflussfaktoren auf Adoption und Rebound mit verschiedenen disziplinären Zugängen untersucht. Eine Integration dieser Zugänge ist oft schwierig, da sie jeweils nur Ausschnitte der Gesamtproblematik beleuchten sowie unterschiedliche Methoden und Maßeinheiten verwenden. Wir stellen hier ein Systemmodell vor, welches das Wirkungsgeflecht zwischen Adoption, Rebound und einer Palette an gemeinsamen Einflussfaktoren kompakt veranschaulicht. Notwendigerweise ist dieses Systemmodell eine Vereinfachung gegenüber fokussierten, disziplinären Zugängen, zeigt aber übergreifende Dynamiken auf. Entlang der Wirkungspfade wird sichtbar, wie sich eine politische Intervention im Geflecht der Einflussfaktoren fortpflanzt und möglicherweise unbeabsichtigte oder kontraproduktive Effekte auslöst.

METHODE

Das Systemmodell (Walker & van Daalen 2013) integriert drei komplementäre Forschungsmethoden die jeweils einen Ausschnitt des Spannungsfeldes Adoption und Nutzung beleuchten: (i) Der semiquantitative Ansatz des Fuzzy Cognitive Mapping übersetzt Expertenwissen über komplexe reale Systeme in einfache Ursache-Wirkung Diagramme (Olazabal 2016). Fuzzy Cognitive Maps bestehen aus Knotenpunkten, die durch Richtungspfeile verbunden sind. Diesen Richtungspfeilen werden „fuzzy values“ (unscharfe Werte) zugewiesen, welche die Einflussstärke darstellen (Özesmi et al. 2004, Papageorgiou et al. 2012). Die Diagramme wurden anhand von 12 ExpertInneninterviews abgeleitet, welche Einflussfaktoren und Akteure für beide Anwendungsbeispiele identifizierten und deren kausale Zusammenhänge gewichteten (Fruhmam et al. 2017).

(ii) Strukturgleichungsmodellierung ist eine Methode aus der multivariaten Statistik zur empirischen Abschätzung eines kausalen Netzwerks von Motiven, die gemeinsam zu einem bestimmten Verbraucherverhalten beitragen (Byrne 2010). Die Motive und ihre Netzwerkzusammenhänge werden aus Theorien zu Umweltverhalten oder zur Technologieakzeptanz abgeleitet (Venkatesh et al. 2003, Steg et al. 2013), statistisch auf Modellpassung getestet und auf die Stärke ihrer gegenseitigen Einflüsse geschätzt. Wir verwenden Umfragedaten von 575 E-Auto-Besitzern und 1455 Personen, die eine Gebäudesanierung durchgeführt haben, um Regressionskoeffizienten für die Auswirkungen von Umweltwerten, sozialen Normen und anderen psychologischen Faktoren auf das Reboundverhalten zu erhalten (Seebauer 2019).

(iii) Angewandte allgemeine Gleichgewichtsmodelle (computable general equilibrium, CGE) berechnen, wie sich ein gezielter Wirtschaftsimpuls über Wirtschaftssektoren ausbreitet und zu endogenen Preis-, Einkommens- und Output-Effekten führt (Allan et al. 2007, Lecca et al. 2014, Broberg et al. 2015). Ein auf die österreichische Wirtschaft kalibriertes CGE-Modell schätzt, wie eine 10%ige Verbesserung der Energieeffizienz in der privaten Verkehrs- und Energienachfrage zu prozentualen Veränderungen des Preises für fossile Brennstoffe, der Nachfrage nach fossilen Brennstoffen, des Bruttoinlandsprodukts (BIP) etc. führt. Des Weiteren wird die Höhe des direkten, indirekten und gesamtwirtschaftlichen Rebound Effekts bestimmt, wenn die tatsächliche Veränderung des Energieverbrauchs schwächer ausgeprägt ist als der anfängliche 10% -Impuls (Kulmer and Seebauer 2019). Aus Platzgründen stellt dieser Artikel die übergreifende Dynamik zwischen Adoption, Rebound und den gemeinsamen Einflussfaktoren in den Mittelpunkt. Details zu den einzelnen Forschungsmethoden sind den jeweiligen Publikationen zu entnehmen (Kulmer und Seebauer 2019, Seebauer 2018, Fruhmam et al. 2017). Die Analysen beziehen sich auf die Fallbeispiele E-Auto und Gebäudesanierung in Österreich.

Die Entwicklung und Anwendung des Systemmodells erfolgt in fünf Schritten: Die (1) Identifikation bestimmt aus jedem der drei (semi)quantitativen Methodenzugänge kritische Elemente, welche die Adoption und die Nutzung von Effizienztechnologien wesentlich beeinflussen. Jeder Methodenzugang beschreibt die Beziehungen zwischen kritischen Elementen hinsichtlich ihrer Kausalrichtung und Einflussstärke. Der Klarheit halber werden Elemente mit geringem oder teilweisem Einfluss weggelassen, wie etwa von Experten als vernachlässigbar eingeschätzte Kausalzusammenhänge und Regressionskoeffizienten mit einem statistischen Signifikanzniveau von $p > 0,05$; der Verbrauch von Nicht-Energie-Gütern in den Haushalten ist zum größten Teil einkommensabhängig und wird im zugrundeliegenden CGE-Modell über den indirekten Rebound Effekt abgedeckt.

Elemente, die in den Systemmodellen beider Fallbeispiele vorkommen	
Element	Definition
Kauf / Investition	E-Auto: Anzahl an E-Auto Käufen; Anteil von E-Autos an Zulassungszahlen Sanierung: Anzahl an Sanierungen; Anzahl an Förderansuchen hinsichtlich teilweiser oder umfassender Sanierung
Direkter Rebound	Nach Adoption der energieeffizienten Technologie steigt die Nachfrage nach der jeweiligen Energiedienstleistung an
Indirekter Rebound	Aufgrund freigewordenen Einkommens oder Moral Licensing werden andere Energiedienstleistungen und Güter vermehrt nachgefragt
Gesamtwirtschaftlicher Rebound	Produktion und Nachfrage verlagern sich in energie- und CO2-intensivere Sektoren
Anschaffungskosten/ Ausgaben	E-Auto: Anschaffungskosten Sanierung: Ausgaben je nach Intensität und Qualität der Sanierung
BIP - Bruttoinlandsprodukt	Bruttowertschöpfung der Volkswirtschaft zuzüglich des Saldos von Gütersteuern und Gütersubventionen
Einsparung CO2-Emissionen	Einsparung an gesamten nationalen Treibhausgasemissionen über alle Wirtschaftssektoren
Massenmedien	Kommunizierte Produktinformationen über die Technologie, CO2-Bilanz, Nachhaltigkeit, Vermittlung der Vorteile diverser Energieeffizienzmaßnahmen
Soziale Norm	Erwartungen des sozialen Netzwerks, ob man sich die energieeffiziente Technologie anschaffen soll
Umweltwerte	Persönliche Überzeugung, dass man einen Beitrag zum Umweltschutz leisten will
Variable Kosten	E-Auto: Kosten pro km, sowohl monetäre als auch Bequemlichkeits-/Zeitkosten Sanierung: Kosten pro behaglich beheizten m ² Wohnfläche nach der Sanierung
Verfügbares Einkommen	Zur Verfügung stehendes Haushaltsbudget; Konsummöglichkeiten werden bis zum Erreichen der persönlichen Sparquote ausgeschöpft
Wissen über Produktmerkmale	E-Auto: Wissen über Reichweite, Info aus Typenschein, CO2 Emissionen pro km, etc. Sanierung: Wissen über Sanierungsmöglichkeiten, Dämmmaterialien, Technologien, etc.
Wohlfahrt	Nutzen aller Haushalte in der Volkswirtschaft, gemessen als Umfang der Konsummöglichkeiten beschränkt durch das verfügbare Einkommen
Zusätzliche Elemente im Systemmodell E-Auto	
Element	Definition
Angebotsqualität: Alternativen zu MIV	Liniennetz und Fahrplandichte des öffentlichen Verkehrs, Radwegenetz
AutohändlerInnen	Kommunizieren Produktinformationen, Kosten, Verbrauch etc.; durch Vertrauensverhältnis sind AutohändlerInnen eine glaubwürdige Informationsquelle
Ladeinfrastruktur	Angebot an öffentlich zugänglichen Ladestationen, der Möglichkeit das E-Auto zu Hause, am Arbeitsplatz oder während Freizeitaktivitäten zu laden, das Angebot an Schnellladestationen, etc.
Produktpalette E-Auto	Angebotsfülle an verschiedenen E-Auto Modellen / E-Auto Klassen
Vorzeigeprojekte	E-Mobilitäts-Modellregionen oder E-Car Sharing Modelle als Kommunikationskanal zur Verbreitung von Produktinformationen; als Plattformen zum Ausprobieren und Testen von E-Autos
Zusätzliche Elemente im Systemmodell Gebäudesanierung	
Element	Definition
Einwandfreie technische Umsetzung	Umsetzung ohne technische/bauliche Mängel, Integration und richtige Dimensionierung von Dämmelementen innerhalb des Gebäudes, Voreinstellungen von Heizsystemen, etc.
Energiearmut	Vor der Sanierung konnte man es sich nicht leisten die Wohnung angemessen warm zu halten
Energieberatung	Kommunikation von Produktinformationen über Gebäudedämmung, Dämmmaterialien und Technologien sowie technologiegerechte Nutzung
Gelegenheitsfenster durch Lebensdauer von alten Technologien	Zeitliche Nähe zum Ende der Lebensdauer sowie Dringlichkeit der Erneuerung von Gebäudeelementen wie z.B. Heizung, Fenster
Gewohnheiten	Automatisiertes Beibehalten von Alltagsroutinen
Installateure & Bauunternehmen	Verantwortlich für Planung, Umsetzung von Sanierungen; agieren über Baumessen, Ausstellungen, etc.; kommunizieren wichtige Produktinformationen (Kosten, Testberichte, etc.).
Komplexität der Antragsstellung (Subventionen)	Notwendige Schritte um Subventionen zu erhalten, Anzahl an involvierten Stellen und Akteuren bis zur Förderzusage(z.B. Banken, Behörden auf Gemeinde-/Landes-/Bundesebene)
Wissen über Verwendung der Technologie	Wissen über richtiges Heizverhalten und Lüften

Tabelle 1: Liste der Elemente je Systemmodell

Die (2) Konsolidierung integriert die Elemente und deren Beziehungen in einem gemeinsamen Rahmen. Elemente, die in mehreren Methodenzugängen eine kritische Rolle einnehmen, werden unter Klärung eines gemeinsamen Grundverständnisses zu einer Einheit verknüpft. Hier kommt die Stärke des Systemmodells zu tragen, dass nahezu alles als Systemelement beschrieben werden kann – Akteure ebenso wie Personenattribute, technische oder infrastrukturelle Rahmenbedingungen. Die konsolidierten Elemente (siehe Tabelle 1) bilden die Basis des Systemmodells.

In der (3) Skalierung werden die disziplinär unterschiedlichen quantitativen Effektgrößen aufeinander abgestimmt. Während sowohl die expertenbasierte Gewichtung des fuzzy cognitive mapping als auch die Regressionskoeffizienten der Befragung die größtmögliche Einflussstärke als perfekten linearen Zusammenhang verstehen, berücksichtigt das makroökonomische CGE Modell, dass zwischen zwei Elementen auch ein exponentieller Zusammenhang bestehen kann. Daher werden alle disziplinären Effektgrößen auf eine gemeinsame Skala von 0 „kein Zusammenhang“ bis 1 „nicht-linearer starker Zusammenhang“ transformiert (siehe Tabelle 2). Die Bestimmung der Effektgrößen erfolgt, indem die Ergebnisse der drei Methoden unter den drei AutorInnen diskutiert werden und ein gemeinsamer Wert vereinbart wird. Diskussion ist ein etablierter Ansatz in der qualitativen Sozialforschung, um Inter-Subjektivität zu erreichen und die Übereinstimmung zwischen verschiedenen Beurteilereinschätzungen zu verbessern (Bortz und Döring 2006). Im Diskussionsprozess werden die theoretischen Grundlagen und empirischen Erkenntnisse jeder Methode erörtert, das Vertrauensniveau der jeweiligen Ergebnisse berücksichtigt, Widersprüche und Unklarheiten geklärt und schlussendlich wird eine gemeinsame Effektgröße abgeleitet, welche diese Überlegungen widerspiegelt (ähnlich dem iterativ-induktiven Prozess der Grounded Theory (Strauss und Corbin 1988) oder bei der Erstellung von IPCC-Berichten (IPCC 2008)).

Effektgröße	Definition
0	Kein Zusammenhang, eine Änderung des einen Elements bewirkt keine Veränderung des anderen Elements
0.25	Mittlerer Zusammenhang, eine Änderung des einen Elements bewirkt eine mittelstarke Änderung des anderen Elements
0.5	Starker Zusammenhang, eine Änderung des einen Elements bewirkt eine starke Änderung des anderen Elements
0.75	Perfekter linearer Zusammenhang, eine Änderung des einen Elements bewirkt eine gleichwertige Änderung des anderen Elements
1	Nicht-linearer Zusammenhang, eine Änderung des einen Elements bewirkt eine vielfache Änderung des anderen Elements
positiver Koeffizient	positive kausale Richtung, eine Erhöhung des einen Elements führt zu einer Erhöhung des anderen Elements
negativer Koeffizient	negative kausale Richtung, eine Erhöhung des einen Elements führt zu einer Verringerung des anderen Elements

Tabelle 2: Skala der Effektgrößen im Systemmodell

Die (4) Integration erstellt das eigentliche Systemmodell. Die Abfolge Adoption, Nutzung (betrachtet als direkter und indirekter Rebound-Effekt) und der aus der Nutzung resultierende kumulierte Effekt (gesamtwirtschaftlicher Rebound-Effekt) bildet das zentrale Analysegerüst. Die Elemente werden hinsichtlich ihrer Wirkung in dieses Gerüst eingereiht und mit gerichteten Pfeilen, die Stärke und Richtung angeben, miteinander verbunden. Das Systemmodell zeigt, welche Elemente wo und wie stark im Spannungsfeld zwischen Adoption und Nutzung agieren und so Rebound-Effekte auslösen oder vermeiden können. Daraus werden Wirkungspfade abgeleitet, die sich entweder gegenseitig verstärken und zu einem großen Effekt aufschaukeln oder sich gegenseitig aufheben und somit den Effekt neutralisieren. Die hier vorgestellten Systemmodelle thematisieren gemeinsame Einflussfaktoren auf Adoption und Nutzung; der offensichtliche direkte Einfluss von Kauf/Investition auf gesamtwirtschaftliche Elemente wird ausgeklammert.

Die (5) Politikanalyse nützt das Systemmodell um Transformationshebel zur Reboundvermeidung zu identifizieren. Ausgewählte Politiken verändern die Ausprägung einzelner Systemelemente. Dieser Impuls pflanzt sich über die Wirkungsbeziehungen zwischen den Elementen, je nach Einflussrichtung und -stärke, durch das System fort. Damit ermöglicht das Systemmodell eine Analyse, welche politischen Maßnahmen oder Maßnahmenbündel am besten geeignet sind, um Rebound-Effekten ganzheitlich entgegenzusteuern und das Spannungsfeld zwischen Adoption und Nutzung aufzulösen. Natürlich ist dieses Systemmodell eine Vereinfachung gegenüber fokussierten, disziplinären Zugängen, zeigt aber übergreifende Dynamiken auf. Eine ausführliche Diskussion und Reflexion im Vergleich zu anderen integrativen Zugängen geben Seebauer et al. (2019).

ERGEBNISSE

Dynamik und Wirkungsverflechtungen der Systemelemente

Das Systemmodell E-Auto umfasst zahlreiche Faktoren, die den Kauf eines E-Autos fördern (Abbildung 1): hohe soziale Norm, hohes verfügbares Einkommen, gut ausgebaute Ladeinfrastruktur, eine breite Produktpalette an verschiedenen E-Fahrzeugen, etc. Neben diesen psychologischen, sozioökonomischen und technischen Faktoren wirken auch AutohändlerInnen und Vorzeigeprojekte als Marktakteure positiv auf die Anschaffung eines E-Autos. Diesen begünstigenden Elementen stehen Anschaffungskosten und Kosten pro Kilometer als hemmende Faktoren gegenüber. Kosten pro Kilometer, verfügbares Einkommen und soziale Norm haben divergente Wirkungen auf Adoption und Nutzung: Einerseits macht die Aussicht auf höhere Kosten pro Kilometer, verursacht etwa durch steigende Strompreise, die Anschaffung eines E-Autos weniger wahrscheinlich; andererseits dämpfen höhere Kosten die Nutzungshäufigkeit und damit den direkten Rebound. Der Vergleich der beiden Effektgrößen zeigt, dass man aus umweltpolitischer Sicht die geringe Einschränkung der Anschaffungsrate durch höhere Kosten pro Kilometer (-0.1) in Kauf nehmen könnte, um im Gegenzug von einer starken Reduktion des direkten Rebounds (0.8) zu profitieren. Verfügbares Einkommen und soziale Norm begünstigen den Kauf eines E-Autos. Haushalte mit höherem verfügbarem Einkommen weisen einen höheren indirekten Rebound auf, sprich einen Mehrkonsum in anderen Bereichen.

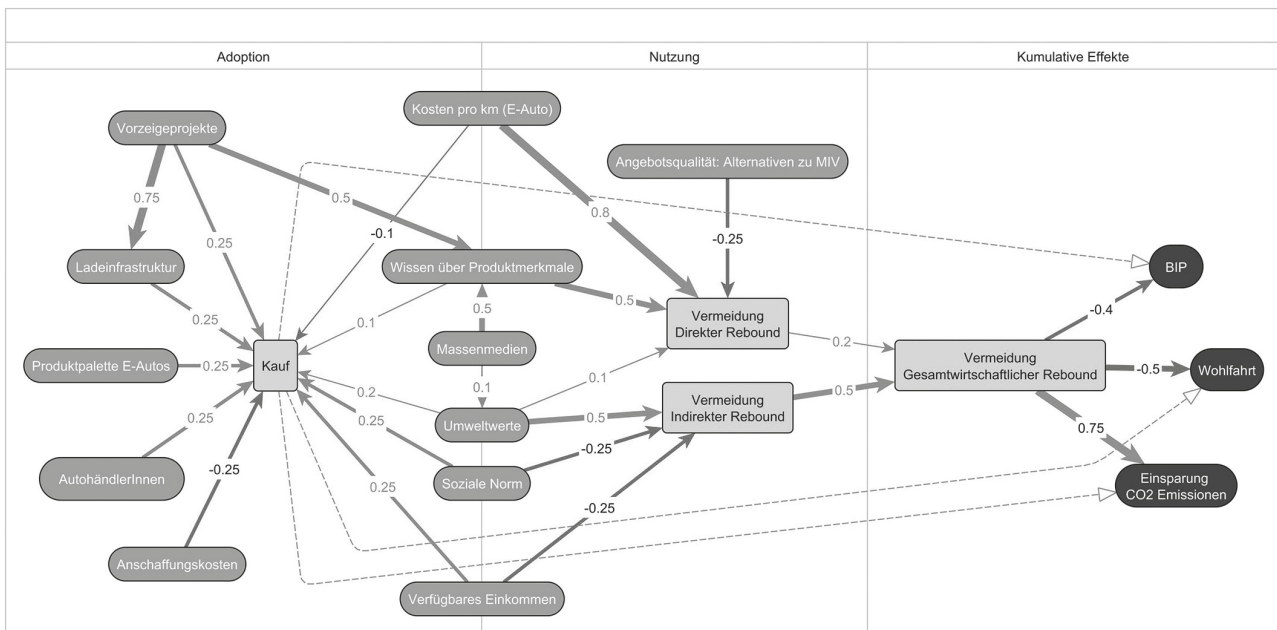


Abbildung 1: Systemmodell E-Auto

Das Systemmodell enthält ebenso Elemente mit konvergenter Wirkungsweise. Wissen über Produktmerkmale und Umweltwerte begünstigen die Anschaffung und vermeiden direkten und indirekten Rebound. Zusammengenommen verdeutlichen die divergenten und konvergenten Effekte von Elementen, dass politische Maßnahmen, die stark auf finanzielle Anreize setzen oder Prestigegewinn in den Vordergrund stellen, kontraproduktive Wirkungen haben. Informative oder bewusstseinsbildende Maßnahmen hingegen lenken sowohl Adoption als auch Nutzung in eine umweltpolitisch erwünschte Richtung.

Das Systemmodell Gebäudesanierung enthält mehr und komplexer vernetzte Elemente (Abbildung 2). Dennoch sind zentrale Wirkungsmuster gleich: Umweltwerte tragen zur Auflösung des Spannungsfeldes Adoption und Nutzung bei, während verfügbares Einkommen, soziale Norm und variable Heizkosten das Spannungsfeld verstärken. Wichtige Einflussfaktoren um Sanierungen anzustoßen sind das Aufkommen eines Gelegenheitsfensters und eine niedrige Komplexität der Antragsstellung für eine öffentliche Subvention. Letztere ist die mit Abstand größte Hürde zur Durchführung einer Sanierung. Bei der Vermeidung des direkten Rebounds kommen im Vergleich zum Systemmodell E-Auto zwei relevante Einflussfaktoren dazu: Energiearmut und technische Umsetzung.

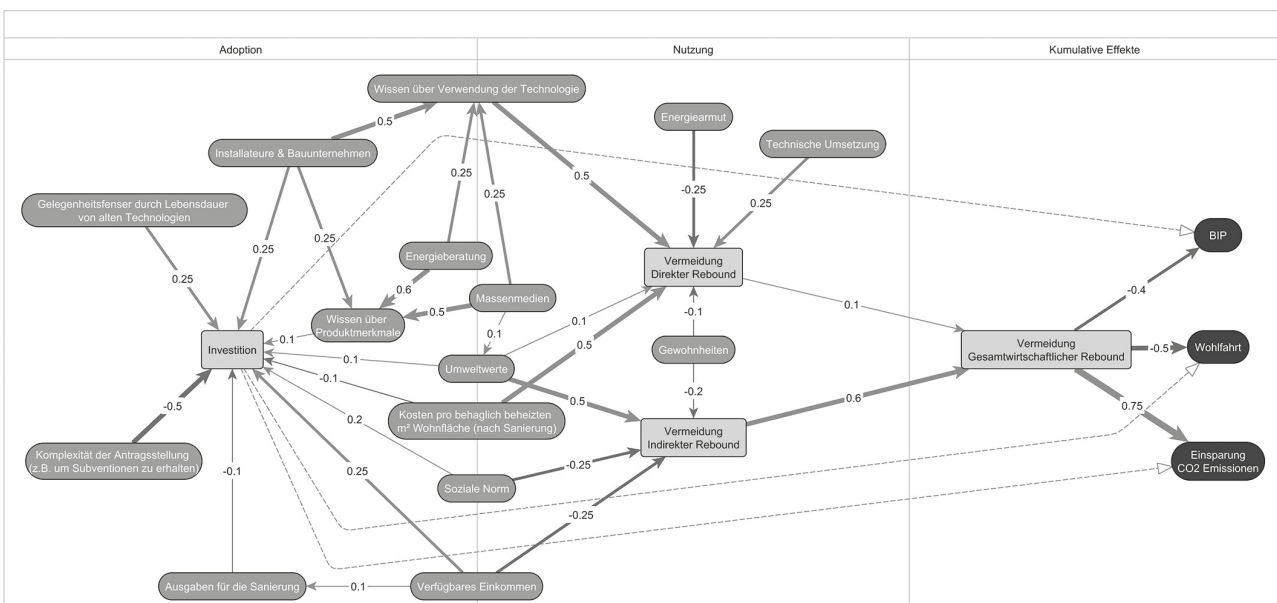


Abbildung 2: Systemmodell Gebäudesanierung

Das Systemmodell Gebäudesanierung verdeutlicht die Rolle von vorgelagerten Akteuren. Massenmedien, Bauunternehmen & Installateure und Energieberatung haben kaum unmittelbaren Einfluss auf Adoption oder Nutzung, wirken jedoch indirekt über andere Systemelemente wie Wissen über Produktmerkmale oder Wissen über Verwendung der Technologie. Letzteres hat hohes Potential direkten Rebound zu vermeiden und wird von allen diesen vorgelagerten Akteuren begünstigt. Diese verzweigten Wirkungskanäle unterstreichen, dass effektive politische Maßnahmen einen Mix aus Elementen ansprechen sollten, anstatt sich nur auf jene Elemente zu konzentrieren, die in einer unmittelbaren Beziehung zu Adoption oder Nutzung stehen.

Beide Systemmodelle weisen auf Schlüsselakteure hin, die widersprüchliche Anreize für Adoption und Nutzung auflösen können. Je mehr Bauunternehmen & Installateuren vertraut wird und je mehr Kompetenz ihnen zugeschrieben wird, desto eher entschließt sich ein Haushalt zur Gebäudesanierung. Gleichzeitig steigern Bauunternehmen & Installateure das Wissen der Haushalte über richtiges Heizen und Lüften und wirken somit dem direkten Rebound entgegen. Beim E-Auto übernehmen Vorzeigeprojekte diese ausbalancierende Rolle. Vorzeigeprojekte schaffen günstige Rahmenbedingungen für den Kauf eines E-Autos und tragen gleichzeitig über Wissensvermittlung zu einer umweltfreundlichen Nutzung bei. Integrative Politikstrategien sind gut beraten, solche Schlüsselakteure gezielt einzubinden.

Der gesamtwirtschaftliche Rebound wird in beiden Fallbeispielen vor allem vom indirekten Rebound angetrieben. Die Verlagerung der Einsparung aus dem Effizienzgewinn in andere Konsumbereiche wirkt um ein Vielfaches stärker als der Anstieg der direkten Nutzung. Ein hoher gesamtwirtschaftlicher Rebound-Effekt spiegelt widersprüchliche Wirtschafts-, Sozial- und Umweltziele wieder. Mehrkonsum führt zu Wirtschaftswachstum und folglich steigen Bruttoinlandsprodukt und, bedingt durch höhere Löhne und mehr Konsummöglichkeiten, Wohlfahrt. Durch den Mehrkonsum wird aber weniger Energie und somit weniger CO₂ als ursprünglich erwartet eingespart. Aufgrund von Wirtschaftsverflechtungen kann dieser Kompensationseffekt so weit gehen, dass die gesamte Energie- und CO₂-Ersparnis durch Mehrkonsum ausgeglichen wird.

Die Systemmodelle beider Fallbeispiele zeigen dieselben Angelpunkte zur Vermeidung von indirektem (und in Folge gesamtwirtschaftlichem) Rebound: (i) hohe Umweltwerte, wodurch der Konsum vorwiegend auf nicht-energieintensive nachhaltige Produkte verlagert wird, (ii) niedrige soziale Norm, da dann der Prestigegewinn durch die energieeffiziente Technologie weniger als Rechtfertigung für Mehrkonsum herangezogen wird, und (iii) niedriges verfügbares Einkommen, wodurch der Konsum verringert oder zu weniger energieintensiven Produkten verlagert wird. Den dritten Aspekt durch ein pauschales Wegbesteuern von Effizienzgewinnen politisch zu nutzen, dürfte jedoch an die Grenzen von Machbarkeit und Akzeptanz stoßen.

POLITIKSIMULATION MITTELS SYSTEMMODELL

Die Politiksimulation skizziert, welche politischen Maßnahmen am besten geeignet sind, um Rebound-Effekten ganzheitlich entgegenzusteuern und das Spannungsfeld Adoption und Nutzung aufzulösen. Die Einflussstärken der Systemelemente und damit die simulierten Politikwirkungen sind naturgemäß mit Unsicherheiten verbunden. In anderen Ländern oder mit einem anderen Strom-Mix für E-Autos, würden

andere Einflussstärken zu anderen numerischen Ergebnissen führen. Hier stehen jedoch die relativen Verhältnisse und Größenordnungen der unterschiedlichen Politikinstrumente im Vordergrund, um zentrale Hebelelemente zu identifizieren. Mit dieser Anwendung erweitern wir gegenwärtige Rebound-Vermeidungspfade (Font Vivanco et al. 2016) um eine vergleichende Wirkungsabschätzung verschiedener politischer Maßnahmen.

Der derzeitige politische Rahmen setzt vorrangig auf fiskalpolitische Instrumente zur Reduktion der variablen Kosten um energieeffiziente Technologien zu forcieren. Zum Beispiel ist in Österreich das E-Auto von der Normverbrauchsabgabe befreit; Strom für den Betrieb des Fahrzeugs wird nicht gesondert besteuert; in manchen Großstädten sind E-Autos von Parkgebühren ausgenommen. Im Systemmodell E-Auto führt eine Reduktion der variablen Kosten pro km um eine Einheit dazu, dass die Anschaffung um 0.1 Einheiten zunimmt, aber zugleich die Nutzung bzw. der direkte Rebound um 0.8 Einheiten steigt. Der direkte Rebound pflanzt sich im System fort und führt zu einer Verringerung der Einsparungen von CO₂-Emissionen in Höhe von 0.12 Einheiten ($0.8 \cdot 0.2 \cdot 0.75$). Ein ähnliches, aber abgeschwächtes Bild zeigt sich im Systemmodell Gebäudesanierung: Geringere Energiesteuern für bestimmte Heizungsenergieträger (z.B. Biomasse, Fernwärme) reduzieren die Kosten pro behaglich beheizten m² Wohnfläche. Das bewirkt eine marginal höhere Investitionstätigkeit von 0.1, aber verringert die CO₂-Einsparungen um 0.0375 ($0.5 \cdot 0.1 \cdot 0.75$).

Bewusstseinsbildende Kampagnen zur Steigerung von Umweltwerten versprechen hingegen deutlich höhere Wirkungen. Die unmittelbare Wirkung auf die Adoption ist auch hier gering (0.2 bei E-Auto, 0.1 bei Gebäudesanierung). Da aber Umweltwerte sowohl den direkten als auch den indirekten Wirkungskanal ansprechen, summieren sich die Wirkungen auf die Einsparung von CO₂-Emissionen. Eine Erhöhung der Umweltwerte um 1 Einheit reduziert bei Gebäudesanierung die CO₂-Emissionen um 0.2325 Einheiten; vor allem über den Wirkungspfad des indirekten Rebounds ($0.5 \cdot 0.6 \cdot 0.75 = 0.225$) und geringfügig über den direkten Rebound ($0.1 \cdot 0.1 \cdot 0.75 = 0.0075$). Beim E-Auto zeigt sich das gleiche Bild ($0.1 \cdot 0.2 \cdot 0.75$ (direkt) + $0.5 \cdot 0.5 \cdot 0.75$ (indirekt) = 0.2025). Die hohe Wirksamkeit macht eine Steigerung von Umweltwerten zu einer attraktiven Interventionsstrategie. Allerdings dürfte eine weitreichende Veränderung von Umweltwerten nur durch langwierigen gesellschaftlichen Wandel zu erreichen sein; finanzielle Anreizinstrumente sind hingegen deutlich einfacher und schneller umzusetzen.

Schlüsselakteure bieten einen rascheren und effektiven Zugang zur Steigerung der Adoptionsrate. Beim E-Auto fördert ein Ausbau von Vorzeigeprojekten um 1 Einheit die Anschaffung eines E-Autos um gesamt 0.4875 Einheiten (direkter Einfluss: 0.25; via Ladeinfrastruktur: $0.75 \cdot 0.25$; via Wissen über Produktmerkmale: $0.5 \cdot 0.1$). Der Einfluss von Vorzeigeprojekten auf Wissen über Produktmerkmale pflanzt sich im System fort und trägt geringfügig zur Einsparung von CO₂-Emissionen mit 0.0375 ($0.5 \cdot 0.5 \cdot 0.2 \cdot 0.75$) bei. Bauunternehmen & Installateure, das Pendant im Systemmodell Gebäudesanierung, weisen ähnliche Wirkungen auf. Diese Akteure erhöhen die Adoption um gesamt 0.275 (direkter Einfluss: 0.25; via Wissen über Produktmerkmale: $0.25 \cdot 0.1$) und führen zu Einsparungen von CO₂-Emissionen um 0.01875 ($0.5 \cdot 0.5 \cdot 0.1 \cdot 0.75$). Die gezielte Einbindung von Schlüsselakteuren bietet sich als flankierende Maßnahme in einem Politikbündel an. Einerseits haben sie einen klaren Haupteffekt auf die Adoption; andererseits generieren sie durch Wissensvermittlung über Produktmerkmale den positiven Nebeneffekt, die CO₂-Emissionen abzuschwächen.

SCHLUSSFOLGERUNGEN UND AUSBLICK

Die breite Markteinführung energieeffizienter Technologien, eine zentrale Strategie in der kohlenstoffarmen Transformation der Gesellschaft, steht im Spannungsfeld von Adoption und Nutzung. Die hier vorgestellten Systemmodelle für die Beispiele E-Auto und Gebäudesanierung integrieren Ergebnisse aus drei komplementären Forschungsmethoden und bestimmen Transformationshebel zur Auflösung dieses Spannungsfeldes. Die Systemmodelle zeigen die Dynamik zwischen Adoption, Rebound-Effekten und der Palette an gemeinsamen Einflussfaktoren.

Finanzielle Anreizinstrumente wie eine Steuerbefreiung von E-Autos (wie in Österreich und Deutschland) oder Ausnahmeregelungen für E-Autos bei städtischen Mautsystemen (wie in London und Stockholm) stehen mitten im Spannungsfeld von Adoption und Nutzung. Sie fördern zwar schwach die Anschaffung der energieeffizienten Technologie, wirken aber gleichzeitig stark auf direkten Rebound. Da finanzielle Anreize technologiespezifisch gesetzt werden und den entscheidenden Wirkungskanal des indirekten Rebounds außer Acht lassen, ist die Gesamtwirkung auf CO₂-Einsparungen gering. Dies steht in klarem Widerspruch zur politischen Beliebtheit marktwirtschaftlich orientierter Instrumente.

Stattdessen sind Maßnahmenbündel zu empfehlen, die mit flankierenden Maßnahmen die Bedeutung von Umweltwerten und Schlüsselakteuren aufgreifen. Diese beschleunigen einerseits die Marktdurchdringung ohne unerwünschte Nebenwirkungen und federn andererseits Wirkungen auf den indirekten Rebound ab. Kampagnen in Massenmedien und sozialen Netzwerken könnten Umweltwerte vermitteln sowie bestehende umweltfreundliche Werthaltungen bestärken. Bauunternehmen & Installateure könnten gezielt geschult werden. Eigene staatliche Förderprogramme könnten die Umsetzung von Vorzeigeprojekten und Testregionen unterstützen.

Solche Maßnahmenbündel können auch ausgleichen, dass verschiedene Maßnahmen ihre Wirkungen über verschiedene Zeiträume entfalten. Steuern und Subventionen wirken unmittelbar und kurzfristig; Schulungen, Kampagnen oder Medienaktivitäten wirken mittelfristig; ein Wertewandel oder eine Änderung von Konsumpraktiken entwickeln sich langfristig über einen Zeitraum von bis zu einer Generation. Die Transformation zu einer kohlenstoffarmen Gesellschaft wird sich über mehrere Jahrzehnte erstrecken. Es scheint daher sinnvoll, kurz-, mittel- und langfristige Maßnahmen systematisch zu staffeln. Für eine zeitliche Dimension müssten die hier vorgestellten Systemmodelle erweitert werden, etwa um Rückkopplungseffekte abzubilden oder um zu berücksichtigen, dass sich über die Zeit Einflussstärken und -richtungen verändern oder neue Systemelemente hinzukommen könnten.

Beide Systemmodelle unterstreichen die dominante Rolle des indirekten Rebound-Effekts. Die Verlagerung der Einsparungen aus dem Effizienzgewinn in andere Konsumbereiche übertrifft bei weitem den Anstieg der direkten Nutzung. Indirekte Verlagerungen münden häufig in eine ansteigende Nutzung von energieintensiven Gütern und Dienstleistungen wie elektronische Geräte, Tourismus und Verkehr. Produktstandards und individuelle Emissionsobergrenzen könnten indirekten Verlagerungen gegensteuern. Produktstandards gibt es bereits in zahlreichen Konsumbereichen (z.B. Emissionsstandards bei Fahrzeugen, Energieeffizienz-Labels bei Elektronikgeräten, diverse ISO-Normen). In je mehr Konsumbereichen Emissionsstandards gelten, desto geringer wäre der Anteil fossiler

Energieträger in der Herstellung von Produkten und Dienstleistungen; folglich wären Verlagerungseffekte in andere Konsumbereiche weniger CO₂ intensiv. Individuelle Emissionsobergrenzen, etwa als jährliches CO₂-Budget pro Person, lenken Konsumverlagerungen in Richtung kohlenstoffarme Produkte und Dienstleistungen und zeigen den Weg hin zu suffizienten Lebensstilen. Im Gegensatz zu Produktstandards ist aber nur schwer vorstellbar, wie individuelle Emissionsobergrenzen in der Praxis administriert und kontrolliert werden könnten.

Um die Wirkungsmechanismen von Produktstandards und individuellen Emissionsobergrenzen abbilden zu könnten, müssten die Systemmodelle beider Systemmodelle erweitert werden. Gegenwärtig sind die Systemmodelle auf die konkreten Effizienztechnologien zugeschnitten und umfassen lediglich jene Elemente, die mittelbar oder unmittelbar auf Adoption und Nutzung wirken. Andere Konsumbereiche, die Produktionsseite der Wirtschaft oder Faktoren weit außerhalb des Entscheidungsspielraums der Haushalte, sind (noch) nicht enthalten. Die Einführung zusätzlicher Systemelemente würde aber das Systemmodell drastisch komplexer machen und die Interpretierbarkeit erschweren.

Der bewusst vereinfachende Zugang dieser Studie reicht hingegen bereits aus, um den grundlegenden Widerspruch ökonomischer Instrumente aufzuzeigen, die gemeinsam mit der Adoption energieeffizienter Technologien auch den Rebound-Effekt in deren anschließender Nutzung fördern. Integrierte Strategien für die Transformation zu einer kohlenstoffarmen Gesellschaft sollten daher die Rolle von sozialpsychologischen Merkmalen und Schlüsselakteuren nicht unterschätzen.

DANKSAGUNG

Das Projekt CATCH (<https://catch.joanneum.at>) wurde aus Mitteln des Klima- und Energiefonds gefördert und im Rahmen der siebten Ausschreibung des Austrian Climate Research Programme durchgeführt (Nr. B464789).

LITERATURVERZEICHNIS

- Allan G, Hanley N, McGregor P, Swales K, Turner K** (2007) The impact of increased efficiency in the industrial use of energy: a computable general equilibrium analysis for the United Kingdom. *Energy Economics* 29(4):779–798. <https://doi.org/10.1016/j.eneco.2006.12.006>
- Azevedo, I.** (2014). Consumer end-use energy efficiency and rebound effects. *Annual Review of Environment and Resources* 39, 393–418.
- Boulanger, P.-M., Couder, J., Marenne, Y., Nemoz, S., Vanhaverbeke, J., Verbruggen, A., Wallenborn, G.** (2013). Household energy consumption and rebound effect. Final report to the Research Programme Science for a Sustainable Development, Brussels: Belgian Science Policy.
- Bortz J, Döring N** (2006) *Forschungsmethoden und Evaluation*. English Translation: Research methods and evaluation. Springer, Berlin
- Broberg T, Berg C, Samakovlis E** (2015) The economy-wide rebound effect from improved energy efficiency in Swedish industries: a general equilibrium analysis. *Energy Policy* 83:26–37. <https://doi.org/10.1016/j.enpol.2015.03.026>

- Byrne BM** (2010) Structural equation modeling with AMOS: basic concepts, applications and programming, 2nd ed. Routledge, New York/London.
- Font Vivanco, D., Kemp, R., van der Voet, E.** (2016). How to deal with the rebound effect? A policy-oriented approach. *Energy Policy* 94, 114–125.
- Friedrichsmeier, T., Matthies, E.** (2015). Rebound effects in energy efficiency – an inefficient debate? *GAIA* 24(2), 80–84.
- Fruhmann, C., Tuerk, A., Kulmer, V., Seebauer, S.** (2017). System complexity as key determinant in achieving efficacious policy transposition and implementation. In: Ashiabor, H. et al. (Hrsg.), *The Green Market Transition. Carbon taxes, energy subsidies and smart instrument mixes. Critical Issues and Environmental Taxation series, vol. 19.* Edward Elgar Publishing, 193–208.
- Geels, F.W.** (2004). From sectoral systems of innovation to socio-technical systems: insights about dynamics and change from sociology and institutional theory. *Research Policy* 33, 897–920.
- Gillingham, K., Rapson, D., Wagner, G.** (2016). The rebound effect and energy efficiency policy. *Review of Environmental Economics and Policy* 10(1), 68–88.
- Guerra, A., Sancho, F.** (2010). Rethinking economy-wide rebound measures: An unbiased proposal. *Energy Efficiency Policies and Strategies with regular papers* 38(11), 6684–94.
- Holtmark B., Skonhoft A.** (2014). The Norwegian support and subsidy policy of electric cars. Should it be adopted by other countries? *Environmental Science & Policy* 42, 160–168.
- IPCC** (2008) Appendix A to the Principles Governing IPCC Work. Procedures for the preparation, review, acceptance, adoption, approval and publication of IPCC reports. <https://www.ipcc.ch/documentation/procedures/>. Accessed 27 Feb 2019.
- Kulmer, V., Seebauer, S.** (2019). How robust are estimates of the rebound effect of energy efficiency improvements? A sensitivity analysis of consumer heterogeneity and elasticities. *Energy Policy* 132, 1–14.
- Lecca P, McGregor PG, Swales JK, Turner K** (2014) The added value from a general equilibrium analysis of increased efficiency in household energy use. *Ecological Economics* 100:51–62. <https://doi.org/10.1016/j.ecolecon.2014.01.008>
- Peters, A., Sonnberger, M., Dütschke, E., Deuschle, J.** (2012). Theoretical perspective on rebound effects from a social science point of view: Working Paper to prepare empirical psychological and sociological studies in the REBOUND project. Working Paper Sustainability and Innovation, S 2/2012, Fraunhofer ISI.
- de la Rue du Can, S., McNeil, M. & Leventis, G.** (2015). Rebound Effects in the Context of Developing Country Efficiency Programs. Final Report. Lawrence Berkeley National Laboratory.
- Olazabal M, Pascual U** (2016) Use of fuzzy cognitive maps to study urban resilience and transformation. *Environment Innovation and Societal Transitions* 18:18–40. <https://doi.org/10.1016/j.eist.2015.06.006>.
- Özesmi U, Özesmi SL** (2004) Ecological models based on people’s knowledge: a multi-step fuzzy cognitive mapping approach. *Ecological Modelling* 176:43–64.
- Papageorgiou EI, Kontogianni A** (2012) Using fuzzy cognitive mapping in environmental decision making and management: a methodological primer and an application. In: Silvern SE, Young SS (eds) *International perspectives on global environmental change.* IntechOpen. <https://doi.org/10.5772/29375>
- Santarius, T.** (2014). Der Rebound-Effekt: Ein blinder Fleck der sozial-ökologischen Gesellschaftstransformation. *GAIA* 23(2), 109–117.
- Seebauer, S.** (2018). The psychology of rebound effects: Explaining energy efficiency rebound behaviours with electric vehicles and building insulation in Austria. *Energy Research & Social Science* 46, 311–320.
- Seebauer, S., Kulmer, V., Fruhmann, C.** (2019). Promoting adoption while avoiding rebound. Integrating disciplinary perspectives on market diffusion and carbon impacts of electric cars and building renovations in Austria. *Energy, Sustainability and Society*, 9(1), 26. doi:10.1186/s13705-019-0212-5.
- Sorrell, S.** (2007). The rebound effect: an assessment of the evidence for economy-wide energy savings from improved energy efficiency. U.K. Energy Research Center Report.
- Steg L, Nordlund A** (2013) Models to explain environmental behavior. In: de Groot JIM, Steg L, van den Berg AE (eds) *Environmental psychology: an introduction.* BPS Blackwell, Chichester.
- Steg, L., Perlaviciute, G., van der Werff, E.** (2015). Understanding the human dimensions of a sustainable energy transition. *Frontiers in Psychology* 6, 805.
- Strauss A, Corbin J** (1998) *Basics of qualitative research: techniques and procedures for developing grounded theory.* Sage, Thousand Oaks CA
- Thomas, B., Azevedo, I.** (2013). Estimating direct and indirect rebound effects for U.S. households with input–output analysis Part 1: Theoretical framework. *Ecological Economics* 86, 199–210.
- Turner, K.** (2009). Negative rebound and disinvestment effects in response to an improvement in energy efficiency in the UK economy. *Energy Economics* 31(5), 648–666.
- Turner, K.** (2013). “Rebound” effects from increased energy efficiency: a time to pause and reflect. *Energy Journal* 34(4), 25–42.
- Venkatesh V, Morris MG, Davis GB, Davis FD** (2003) User acceptance of information technology: toward a unified view. *MIS Quarterly* 27(3):425–478. <https://doi.org/10.2307/30036540>

Walker W., van Daalen C. (2013). System Models for Policy Analysis. In: Thissen W., Walker W. (eds) Public Policy Analysis. International Series in Operations Research & Management Science, vol. 179. Springer, Boston, MA

Wolf, A., Seebauer, S. (2014). Technology adoption of electric bicycles: A survey among early adopters. Transportation Research Part A 69, 196-211.

AUTORINNEN

VERONIKA KULMER

JOANNEUM RESEARCH Forschungsgesellschaft mbH
LIFE Zentrum für Klima, Energie und Gesellschaft
Waagner-Biro-Straße 100/9, 8020 Graz
E: veronika.kulmer@joanneum.at
Tel: +43 316 876 7651

SEBASTIAN SEEBAUER

JOANNEUM RESEARCH Forschungsgesellschaft mbH
LIFE Zentrum für Klima, Energie und Gesellschaft
Waagner-Biro-Straße 100/9, 8020 Graz
E: sebastian.seebauer@joanneum.at

CLAUDIA FRUHMANN

JOANNEUM RESEARCH Forschungsgesellschaft mbH
LIFE Zentrum für Klima, Energie und Gesellschaft
Waagner-Biro-Straße 100/9, 8020 Graz
E: cfruhmann@outlook.com

SCHLAGWORTE:

Reboundeffekt, Systemmodell, Energieeffizienzpolitik

KEYWORDS:

rebound effect, system model, energy efficiency policy

Eine englischsprachige, erweiterte Fassung dieses Artikels ist erschienen als:

Seebauer, S., Kulmer, V., Fruhmann, C. (2019). Promoting adoption while avoiding rebound. Integrating disciplinary perspectives on market diffusion and carbon impacts of electric cars and building renovations in Austria. *Energy, Sustainability and Society*, 9(1), 26. doi:10.1186/s13705-019-0212-5. (CC BY 4.0).

FAIR PROJECTS – BAD DATA?

EVALUATING THE GENDER BALANCE IN SCIENCE PROJECTS

MAGDALENA KLEINBERGER-PIERER, SIMON POHN-WEIDINGER AND KARIN GRASENICK

DOI: 10.22163/ftval.2020.471

ABSTRACT

Researchers are more and more frequently required to report the gender balance of their teams in order to receive funding for their research projects. In Europe for instance, Horizon 2020 guidelines determine that applications with a balanced, 50/50 representation of women and men will be given preference, and each project must at least justify the composition of their teams and leadership positions.

In scientific fields where women are still a minority, like robotics or artificial intelligence, or many STEM fields in general, this approach has left the applicants with the issue of how to justify the ratio of women on their teams. For individual projects and project leaders, realistic objectives are required that take into account specific framework conditions in different scientific fields.

This paper examines approaches to measure the disciplinary background and career development of women and men in science by assessing a range of available data sources. It provides insights on how to derive figures allowing science projects to evaluate their gender ratios against a possible underrepresentation of women.

Acknowledgement: Our research has received funding from the European Union's Horizon 2020 Framework Programme for Research and Innovation under the Specific Grant Agreement No. 785907 (Human Brain Project SGA2).

TABLE OF CONTENT

1	Introduction	60
2	Benchmarking gender-balance in R&D-projects – a data-based approach	61
2.1	Relevant data sources	62
2.2	Selected indicators	63
2.3	Case Study Reference Data for Computer Sciences	66
3	Summary and Conclusions	69
4	References / Literature	70

1. INTRODUCTION

Despite many efforts and gradual increases, women's share in many academic STEM (Science, Technology, Engineering, Mathematics and Medicine) fields is low, although with wide variations at regional and national levels as well as for scientific disciplines and career levels. In the EU, 48% of all doctoral graduates were women in 2016: They were over-represented for instance in education (68%), but clearly under-represented in information and communication technologies (21%) or engineering, manufacturing and construction (29%). At the level of professors (grade A staff), women's shares within the EU are particularly low for engineering and technology (12%) or natural sciences (18%), and highest for the humanities (32%). (All figures from European Commission 2019 - SHE FIGURES 2018).

The reasons for the disproportionately low participation of women in many fields are manifold and have been discussed extensively (see e.g. The Lancet 2019; PLOTINA 2015; Ranga, Gupta & Etzkowitz 2012; National Academy of Sciences et al. 2007): They include societal causes (such as role models, typical male and female fields, negative stereotypes) and early career choices taken in school – which influence the available women's talent pool at the very beginning of academic education. But even when women decide to pursue their studies in STEM fields, there is a high probability that they are thrown off their career path by a phenomenon labelled **"leaky pipeline"**. This can be observed in all fields of science (and in industry as well as in all other societal domains): Along career and hierarchical levels, the representation of women declines continuously with each step up. This is also true for fields where women's shares of bachelor / master students and doctorates have been large for years and thus there would be enough qualified women available for high-level positions (see National Academy of Sciences et al. 2007). The leaky pipeline can be considered a result of the interplay of many hindering factors, such as institutional climate at universities and research organisations and unfavourable working conditions (i.e., short-term contracts, long working hours, lack of family friendliness and work-life-balance). Also, cultural and structural, frequently unconscious biases play a decisive part in this context (e.g., homo-sociality in networks, "all-boys team-networking", male dominance of evaluation and peer-review processes). Unconscious biases influence our perceptions and decisions, based on previous experiences, associations and preferences already made in the past (see e.g. Staats et al. 2017 for an overview). Despite striving for objectivity, we assess a person not only by performance, but also by affiliation to a specific group that we (unconsciously) consider as positive or negative. Also, it can be a question of visibility, for instance when setting up a research or project team. The principal investigators

and proposal writers refer to people they already know, who are part of their network, who have contributed substantially in the past and have been able to attract attention.

At the organizational level, many universities and research organizations act to increase the number of women, to remove discriminating barriers in science and decision-making, to overcome unconscious biases and to counteract the leaky pipeline. These actions include for instance advertising certain fields of science in schools or even kindergarten or offering summer schools to explore what it might be like to study. In terms of keeping qualified women in science and research, gender equality plans are developed and implemented (as for instance supported by the ongoing PLOTINA project – see PLOTINA 2015).

With regards to research funding, European and national funding agencies increasingly require **projects to report on their gender balance** and make this a criterion in their funding decisions. This is an attempt to urge leading scientists to improve the gender balance of their teams. For instance, projects funded by the EU's Research and Innovation programme Horizon 2020 should pay attention to gender equality in terms of human resources (balanced participation of women and men in the research teams), as well as in research contents (European Union 2019). In particular, the criteria for human resources is explained as follows:

“When applying for a grant under Horizon 2020, you are encouraged to promote gender balance at all levels in your teams and in management structures. Applicants should seek at having a balanced participation, as close as possible to 50/50, of both men and women in the teams and among the leading roles. At the evaluation stage, gender balance in staff is one of the ranking factors that come into play to prioritise the proposals above the threshold with same scores. When it is used, evaluators need to compare the shares of men and women in the personnel named in the proposals (in Part B, section 4.1, of the proposal template) and they will rank higher the proposal with the share closer to 50/50.” (European Union 2019).

Besides being a value in its own right, gender equality makes sense in terms of innovation, better results or positive working atmosphere (see e.g. Schiebinger et al. 2011-2018; Bert 2018). And many measures and initiatives are already in place that support this development (see above).

But in the short term, given the current situation as described above, demanding a balanced participation of 50:50 in projects is highly unrealistic for a range of scientific fields, in particular from the STEM domain. Still, the pressure is there, and project leaders are constantly asked to raise the number of women, regardless of the fields of science involved or the aim of the project. Researchers and project applicants confronted with respective criteria might be unsettled by such demands and question whether they actually have the responsibility to act within their project and the possibility to do so given the framework conditions in their specific scientific fields (see as well Grasenick 2019). Thus, for individual projects and project leaders, realistic goals are required. In order to get an idea of which target values could be realistic and at the same time challenging, knowledge of the actual proportion of women in research is a necessary prerequisite for a target-oriented discussion.

In this context, this paper addresses the following questions: How should a research team determine its gender balance goals? What is a good and fair gender balance at different career levels (PhD students, work package leaders etc.) and for different scientific disciplines? How can it be measured? What are good reference numbers? These questions are not only of relevance for research teams, projects or project consortia who apply for funding, but also for the funding agencies, programme managers and evaluators of funding programmes, as well as for universities and research organisations (e.g. in the context of performance agreements and the goals that are determined by them).

The aim of this paper is to make a contribution to an evidence-based discussion in this field. We have analysed international data sources that cover gender-specific data in research and assessed the usefulness of these sources. After providing an overview of these sources and potentially suitable indicators, we apply them to a case study from the field of Computer Sciences. We conclude with our suggestions of how to proceed when assessing and setting goals for a team's gender balance.

2. BENCHMARKING GENDER-BALANCE IN R&D-PROJECTS – A DATA-BASED APPROACH

In the last decades, a series of attempts have been undertaken to establish gender-related data sources in science. At the institutional level, many universities and research organisations have implemented gender-monitoring systems. There are also a range of data compilations that cover and compare many countries and/or organizations.

Before discussing the different data sources in more detail, it must be clarified which criteria a database should fulfil to be relevant for the purpose of a gender-balance benchmark.

CRITERIA FOR THE RELEVANCE OF DATA SOURCES

- **Regular update of data:** Several publications which offer highly relevant insights on the gender balance in sciences and related professional fields are not available on a regular basis. This includes for instance the Elsevier report on gender in the global research landscape 2017 (see below). Some of these reports set a specific disciplinary focus, for instance, the World Economic Forum's "Global Gender Gap Report" 2018 focusses particularly on Artificial Intelligence¹. However, one-time reports or studies are reduced in their value, since the reference values will not be updated on a regular basis. An essential criterion for the data sources we assess here is that they are updated regularly.

In addition to this formal criterion, several content criteria are of relevance. Databases in question should give information about those framework factors which have an impact on the number of women re-

1 See World Economic Forum (2018). Another example is the "Global Survey of Physicists" conducted in 2009/2010 by the Union of Pure and Applied Physics (IUPAP). Currently, the same organisation conducts a global survey and study of publication patterns of Mathematical, Computing and Natural Scientists, see Butcher et al. (2019).

presented in research (see e.g. Ranga, Gupta & Etzkowitz 2012). Some of the relevant dimensions are:

- **Field of research:** Women's participation varies substantially between different research fields. In general, in Humanities and Social Sciences their participation is higher than in most Engineering or Natural Sciences.
- **Countries:** Women's participation is also a question of a country's "research culture". Even within the same field of research, one can find very different participation rates depending on the country.
- **Leaky pipeline:** The so called "leaky pipeline" is a well-known pattern across all sciences and can be summarized as follows: The higher the career-level the lower the representation of women.

So, when searching and selecting a database, it must be clarified which of these parameters are covered by the provided indicators. To anticipate our findings: None of the reports and databases analysed covers all relevant dimension at once. There are well elaborated databases that are certainly useful for specific purposes, but each one of them has restrictions. Some of them are limited to certain countries² or regions³. Others only cover certain research fields⁴ or will not be updated.

In this chapter we provide an overview of databases and selected indicators that we consider useful to assess whether the gender composition of the research team is well balanced within the (international) research framework. To test if the suggested indicators are feasible, we elaborate a case study in the research field of computer sciences. The findings are summarized in a concluding chapter.

2.1 RELEVANT DATA SOURCES

Since our focus is on international research groups, we have limited the search to those databases that provide data on an international level. A screening of international data sources shows that there are at least four approaches which can help research consortia in assessing their gender balance. Three of them provide country-specific data, one is collecting comparable (publication) data on the level of universities ("Leiden ranking").

CROSS-COUNTRY DATABASE

European Commission (2019). "She Figures 2018": Probably the most useful source in this context are the "She Figures", which are updated every three years by the European Commission. The last release of "She Figures 2018" was published in February 2019 (European Commission 2019). The report provides data for a wide range of gender-specific indicators in the field of R&D, universities and innovation. All data are presented by sex and at the individual country level as well as the broader EU level for the current 28 EU Member States and the associated countries. Many indicators are published for six main Fields Of R&D ("FORD") according to the OECD Frascati Manual⁵. Other indicators like "doctoral graduates" are based on the ISCED-F 2013 classification (UNESCO Institute of Statistics, 2014) which distinguishes 29 narrow fields of education and training ("fields of studies -FoS") organised in 10 broad groups⁶. One limitation of this source is that most countries from outside of the EU are not taken into account.

Elsevier (2017). "Gender in the Global Research Landscape. Analysis of research performance through a gender lens across 20 years, 12 geographies, and 27 subject areas": Unlike "She Figures" this report covers worldwide data and shows development over 20 years. However, there is no clear evidence for a regular update of the report. The study is based on publications from Scopus and ScienceDirect and a combination of Author Profiles with gender-name data from different platforms. The report, published in 2017, includes shares and absolute numbers of men and women as researchers by 27 different subject areas⁷. These subject areas are based on the classification which is used by the Scopus database and differs from classifications like FoS or FORD mentioned above. Data are provided for 12 comparator countries: the EU-28 as a group, as well as four selected individual EU member states (United Kingdom, France, Denmark, Portugal) and seven countries from other continents (United States, Canada, Australia, Brazil, Japan, Mexico, Chile).

UNESCO data-base and UNESCO Science Report (2015): The UNESCO Science Report is published every five years and maps STI (Science, Technology & Innovation) data around the world and relies on data provided by the UNESCO database (UNESCO Institute for Statistics 2019). One chapter in the current report (2015) deals with the gender gap in science and engineering and provides general gender related data on an international level (country/world-regions). Some of the indicators are presented by broad science fields ("FORD"). Apart from the report

2 For instance, the "Science in Australia Gender Equity" (SAGE) initiative (SAGE 2019).

3 A well elaborated example is the gender report of the universities of Nordrhein-Westfalen (Germany) that is published every three years. This report, and a well-presented online portal provide detailed gender related data (Netzwerk Frauen- und Geschlechterforschung NRW 2019).

4 For instance, for biomedical research fields there is a very detailed data collection based on worldwide publications (see Holman, Stuart-Fox & Hauser 2018).

5 According to the OECD Frascati Manual (OECD 2015) these fields are (1) Natural sciences (2) Engineering and technology, (3) Medical sciences, (4) Agricultural sciences, (5) Social sciences and (6) Humanities. The breakdown of researchers by field of R&D is based on the field where they work and not according to the field of their qualification (see also European Commission 2019 - She Figures 2018, p. 188).

6 According to the ISCED-F 2013 classification, these 10 broad groups are: (1) education; (2) humanities and arts; (3) social sciences, journalism and information; (4) business administration and law; (5) natural sciences, mathematics and statistics; (6) information and communication technology; (7) engineering, manufacturing and construction; (8) agriculture, forestry, fisheries and veterinary; (9) health and welfare; and (10) services.

7 Scopus is published by Elsevier. The classification distinguishes between 27 areas, which are: Multidisciplinary (journals like Nature and Science), Agricultural & Biological Sciences, Arts & Humanities, Biochemistry/Genetics/Molecular Biology, Business, Management, & Accounting, Chemical Engineering, Chemistry, Computer Science, Decision Sciences, Dentistry, Earth & Planetary Sciences, Economics/Econometrics/Finance, Energy, Engineering, Environmental Science, Health Professions, Immunology & Microbiology, Materials Science, Mathematics, Medicine, Neuroscience, Nursing, Pharmacology, Toxicology, & Pharmaceuticals, Physics & Astronomy, Psychology, Social Sciences, Veterinary

UNESCO is also running an open source database where gender related data on science, technology and innovation (STI) by country and broad fields of sciences are monitored (UNESCO Institute for Statistics 2019, data.uis.unesco.org/). In general, the database covers 163 countries worldwide (including all 12 countries that are used as comparators by Elsevier, for instance, and all EU member states). But country data are incomplete or even fragmentary for important indicators.

CROSS-ORGANISATIONAL DATA BASE

Leiden Ranking (2019): The current CWTS Leiden Ranking 2019 includes information of scientific performance of nearly 1000 major universities from 56 different countries. This covers for instance universities from most EU member states⁸ and from all 12 comparator countries that were used in the Elsevier Report (2017). The CWTS Leiden Ranking benchmarks universities according to their publication performance. The CWTS Leiden Ranking 2019 is based on bibliographic data from the Web of Science database and presents a variety of indicators to explore the performance of universities from different angles. Since 2019 it also includes gender indicators based on publications. Data are provided on the level of universities, which by themselves are mapped by countries. The data are presented along five scientific fields (Biomedical and health sciences, Life and earth sciences, Mathematics and computer science, Physical sciences and engineering, Social sciences and humanities). This classification is more restricted compared to the data of She Figures, since data of narrow scientific fields are not available.

OVERVIEW OF DATABASES

Table 1 gives an overview which relevant dimension are covered by each database:

- All databases do cover different research fields to a certain extent. Some provide data for detailed fields of research and development ("FORD") or fields of study ("FoS"), others are limited to broad classifications.
- Concerning the coverage of countries, all database provide data on the level of countries. However, in the Elsevier and UNESCO databases, countries that are important from a European perspective are missing. The Leiden ranking is by concept comparing research organisations and not countries. But as the organisations are also linked to countries one can indirectly derive country specific information.
- Only She Figures includes different career stage (like doctoral graduates or professors).

database	Database covering different...			
	fields of research	countries	research organisations	career stages
UNESCO Science Report	yes	(yes)	no	no
Elsevier – Gender report	yes	(yes)	no	no
She Figures	yes	yes	no	yes
Leiden ranking	yes	(no)	yes	no

Table 1: Relevant databases and the dimensions they cover
source: own compilation

INTRA-ORGANISATIONAL DATABASES

In the databases discussed above, we find three cross-country approaches and one cross-organisational database ("Leiden ranking"). They provide orientation on the international situation concerning specific gender-related indicators. These data can serve as a benchmark for checking the gender balance of specific research consortia against the international context.

A similar benchmark could also be derived directly in a more bottom-up approach, with data from the partner-organisations of research consortia themselves: Today many universities and research institutions provide annual gender monitoring reports, which give an up-to-date overview of women's participation in the area of teaching and research. Most of these monitoring reports also include women's participation by career stage and science field (which normally has the structure of a "leaky-pipeline"). Even if a university does not publish such reports, internal monitoring systems are usually in place including the respective information. For example, the University of Manchester (UOM) publishes an Equality Information Report⁹ with information on all staff and students differentiated by diversity dimensions. It includes statistics of the proportion of women for the university overall (all sciences aggregated). However, this has to be based on more detailed information regarding departments and science fields. For instance, the Department of Computer Science, reports that "*women make up 24% of academic staff, higher than UK average*" (University of Manchester – Department of Computer Science 2019). Similarly, the Technical University Munich (TUM) offers a diversity reader, which includes an overview on women at different career levels (Technische Universität München 2018).

2.2 SELECTED INDICATORS

Most of the databases we discussed above include a series of gender-related indicators, which cover a wide range of topics. The screening of these indicators reveals that for the purpose of extracting reference values, only a few of them are of relevance. These will be presented in the following:

8 The ranking 2019 includes no universities from the following five EU member states: Bulgaria, Cyprus, Latvia, Luxembourg, Malta.

9 For the 2019 report see University of Manchester (2019). E.g. women make up 49.8% of the workforce overall and 42.2% of academic staff, but no information on the level of departments or scientific disciplines is given.

PROPORTION OF WOMEN RESEARCHERS IN GENERAL AND FOR SPECIFIC RESEARCH FIELDS

“Share of female researchers by country, 2013 or closest year (%)”, (Source: UNESCO Science Report, p. 87f.): Values are given for countries and regions all over the world for all fields of sciences. For (world-)regions like EU, Latin America or West Asia single values are published. Unfortunately for most individual countries, only ranges of values are published. For instance, according to the report, in Austria women’s share of researchers is between 25% and 34,9%. A further restriction is that the indicator does not differentiate between different levels of career stages.

“Female researchers by field of science, 2013 or closest year (%)” (Source: UNESCO Science Report, p. 87): This indicator goes one step further than the former indicator. It covers more than 75 countries from all over the world and differentiates broad research fields (Natural sciences, Engineering and technology, Medical sciences, Agricultural sciences and Social sciences and humanities). Although values are available for more than 75 countries, for many important countries in science there are no data (e.g. USA, Germany, UK, Italy, France, Austria) provided. Furthermore, the indicator also does not differentiate between different levels of career stages.

“Proportion and number of researchers by gender [...] and research areas, 2011 – 2015”, (Source: Elsevier): This indicator is provided in two variations. The first gives an overall picture of women researchers (p.18), the second indicator (p. 24) differentiates between 27 research areas (e.g. Medicine, Engineering, Neuroscience) according to the Scopus classification of fields of study (ASJC). The indicators cover selected countries outside the EU (e.g. US, CAN, BR, AUS) as well as the EU-28 and selected individual EU member-states (UK, FR, P, DK). However, important research-intense countries like GER¹⁰, IT, AT, CH etc. are not included. Therefore, there are certain restrictions concerning the regional coverage of the data. Another disadvantage is that the indicator does not differentiate between career levels.

CAREER LEVEL

Doctoral graduates: “Proportion (%) of women among doctoral graduates”, 2016. (Source: She Figures 2018): The indicator is provided for all researchers (p. 20), as well as by broad fields of study¹¹ (p.23). Furthermore, for selected STEM-fields like natural sciences/mathematics and statistics, ICT and engineering/manufacturing and construction (p.26) data are also available by narrow fields of study¹². Data are presented by country and EU-28. Usually, doctoral graduates are a very important group in academic research teams, as they frequently are responsible for the (time-consuming) field research (laboratory research etc.). Therefore, this indicator can serve as a meaningful benchmark for research groups applying for projects.

Full professorship: proportion (%) of women among Grade A positions (Source: She Figures 2018): For the level of leadership the “She Figures 2018” provides data for full professors by gender. In the terminology of the report, this group is labelled as Grade A¹³: In this context, two indicators are published: The indicator “Evolution of the proportion (%) of women among Grade A positions, 2013 vs. 2016”, (p. 119) provides data for all sciences (aggregated) and a second indicator provides data by six broad fields of R&D (“The proportion (%) of woman among grade A staff among by main field of R&D¹⁴, p. 121).

PUBLICATIONS:

Proportion of women researchers in publications. (Source: Leiden Ranking 2019) (“*The number of female authorships as a proportion of a university’s number of male and female authorships*”)¹⁵: Data are available on the level of universities. However, the Leiden ranking does not differentiate between first and co-authors or the scientific impact of the publications. The proportion of women’s authorships from this ranking could be used as a rough reference (though not a direct benchmark of publications) for the gender balance of a specific research team or project consortium.

10 However, there is a separate report for Germany that was published before the Global Report (Elsevier 2015).

11 The ISCED-F 2013 classification (UNESCO Institute of Statistics, 2014) distinguishes 29 narrow fields of education and training organised in 10 broad groups: education; humanities and arts; social sciences, journalism and information; business administration and law; natural sciences, mathematics and statistics; information and communication technology; engineering, manufacturing and construction; agriculture, forestry, fisheries and veterinary; health and welfare; and services.

12 Selected narrow fields of study: Natural sciences, mathematics and statistics (Biological and related sciences; Environment; Physical sciences; Mathematics and statistics); Information and Communication Technologies; Engineering, manufacturing and construction (Engineering and engineering trades; Manufacturing and processing; Architecture and construction)

13 “Grade A either corresponds to the rank of full professor, or to the highest post at which research is normally conducted.” (SHE Figures 2018, p. 118 and p 190)

14 Fields of R&D: NS = natural sciences; ET = engineering and technology; MS = medical sciences; AS= agricultural sciences; SS = social sciences; H = humanities.

15 This indicator is abbreviated as PA(F)MF in the Leiden Ranking.

Table 2 summarises the discussed indicators.

	Source	Fields of research/study			Country level
		All fields (aggregated)	broad fields	narrow field	
Proportion of women Researchers (general)					
"Share of female researchers, 2013 (%) or closest year (%)"	UNESCO Science report, p. 88ff. (UNESCO open source data)	yes	No	no	yes (but important developed countries missing, for many countries only value ranges available)
"Female researchers by field of science, 2013 or closest year (%)"	UNESCO Science report, p. 87ff. (UNESCO open source data)	no	Yes (FORD*)	no	yes (but important developed countries missing)
Proportion of researchers by gender (2011-2015) and subject area	Elsevier, p. 24	yes	No	Yes (ASJC**)	yes (but only selected EU- and non-EU- states, EU28 included)
Career Level					
Doctoral graduates: Proportion (%) of women among doctoral graduates, 2016	She Figures, p. 19f, p. 23, p. 26	yes	Yes (FoS**)	Yes, but only for STEM- fields (FoS**)	yes
Grade A: proportion (%) of women among Grade A positions, 2016	She Figures, p. 119, p121	yes	Yes	no	Yes (some countries missing)
Publication					
"The number of female authorships as a proportion of a university's number of male and female authorships (2014-2017)"	Leiden Ranking	no	Yes	no	(no) (but indirect by location of universities)

Table 2 Useful gender-related indicators from international data bases / reports

Source: own compilation, *FORD: Field of R&D, **FoS: field of study, ***ASJC: All Science Journal Classification Codes (Scopus)

UNIVERSITY BASED INDICATORS FOR TEAMS AND CONSORTIA

Benchmarking the gender balance of a research team against the data from the above databases is one possible way of estimating the gender balance of a team within the international research landscape. The international perspective is highly relevant, since at least for post-doc positions onward, it is the global scientific community that's relevant and teams at one university or institute are often composed internationally.

However, a project team or consortia might be composed of scientists already working at the specific university or partner organisations involved. In this case, the relevant reference to be considered should be drawn from the statistics available at these universities or consortium partners. Relevant benchmark-indicators could be:

- **Proportion of researchers** by gender in all partner organisations and by broad und narrow fields of R&D.
- **Proportion of doctoral graduates** by gender in all partner organisations and by broad und narrow fields of R&D.

- **Proportion of professors** by gender in all partner organisations and by broad und narrow fields of R&D.

The advantage of such data is that they are very close to the consortia and even if not published, most of the organisations should have gender related data. The values of the indicators (e.g. share of women among doctoral researchers in Biomedicine) can be collected for each university or partner-organisation of the consortia. The values can serve as a benchmark corridor for women's participation within the project team. However, a few challenges arise:

- **Not standardised:** Unlike publications like "She Figures 2018" which provide standardised and comparable data, research organisations collect their data based on different classifications. There is no standardised monitoring system for all universities. Therefore, the indicator values that are collected in two different organisations may express similar things but may be different in detail (e.g. concerning the classifications of field of research, the definition of professorships, PhD students vs. doc-

toral graduations). Due to this lack of standardisation, one must be careful while interpreting the data.

- **Big consortia:** Sometimes, especially if consortia consist of many partner organisations, the compilation of data can be a time-consuming exercise. A possible way to solve this problem could be to benchmark data only from selected (leading) partners.
- **Interdisciplinary consortia:** Problems also arise in interdisciplinary research groups, since the average of women's participation differs significantly between research areas. This issue is not easy to handle. Using average values across a range of disciplines might lead to false conclusions, as particularly high and low values will cancel each other out. In large consortia, each discipline could be analysed separately, whereas in small consortia, it is difficult to derive adequate values. In this case, it could be an option to focus on one core discipline within the project.

These limitations do not mean that collecting data directly from partner organisations is of no use. The contrary is the case. The data give a close-to-project-view of the gender balance of the partner organisations which can be benchmarked against the data of the project team. However, data must be interpreted with caution, due to the mentioned limitations which result from the bottom up data collection.

2.3 CASE STUDY REFERENCE DATA FOR COMPUTER SCIENCES

Having discussed existing data sources and possible indicators, we propose to look at all these indicators, as there is no one perfect solution. Going one step further, we now set out to test if and how this is feasible for a hypothetical research consortium. We have chosen computer science (or ICT – Information and Communication Technology), as a field of science where women are a minority. With the intention to be as realistic as possible when choosing specific organisations for a hypothetical consortium, we considered the “Neurobotics Platform”, which is one of the eight main subprojects of the “Human Brain Project” (HBP)¹⁶. Neurobotics in general is interdisciplinary, including specialist fields of computer sciences and drawing on the connection with brain research to understand and implement aspects of motion and sensation, learning and reasoning. Eight organisations from six countries are listed as partners of this platform¹⁷:

- Ecole Polytechnique Fédérale de Lausanne (EPFL) (CH)
- Technische Universität München (TUM - Robotics and Embedded Systems) (GER)
- Fortiss, an-Institut der Technischen Universität München (Fortiss) (GER)
- Forschungszentrum Informatik am Karlsruher Institut für Technologie (FZI) (GER)

- SSSUP - Istituto di BioRobotica (Scuola Superiore Sant'Anna) (IT)
- Universiteit Gent (BE)
- Universidad de Granada (ES)
- Technical University of Denmark – DTU (DK)

AVAILABLE DATA

For the analysis on a detailed level, we used the field “Computer (and Information) Science” or “Information and Communication Technology”, where available. According to the Frascati Manual (OECD 2015), this is part of Natural Sciences as one of six broad fields of research and development (“FORD”).

The following data sources could be used:

- **Country-level:** For the benchmark on country level data from “She Figures 2018” and partly “Elsevier” were useful. For the latter only data are available for EU-28 for “all sciences” (aggregated) as well as for “Computer Sciences”. Data from the “UNESCO science report” were too unspecific to be helpful (gender related data for the partner-countries are only available in an aggregated way for “all sciences”, data are provided as rather broad ranges, e.g. share of women in all sciences in Germany between 25%-34,9%). Therefore, data from the UNESCO report have not been included in the data compilations.
- **Organisational level:** For this paper, we only conducted desk-research of online available data and did not contact every research organisation directly asking for gender related data. Desk research showed that only two partner organisations (EPFL, TUM) have published gender-related data online. The partner organisation FORTISS as a so called “an-Institut” can in this respect be considered as part of the TUM. Therefore, data are available for three organisations. The definition of the published indicators varies between the organisations. For example, the TUM publishes “completed promotions” while for EPFL the “number of PhD-Students” are published. There are also differences in the classification of the research fields. While data from TUM focus on “Informatik”, data from EPFL cover the fields of “Computer and Communication Sciences”. Therefore, data allow only a rough benchmark.
- **Publication:** In the Leiden-Ranking six of the eight partner organisations are included.

The benchmark for women's participation reveals the following picture.

SHARE OF WOMEN RESEARCHERS IN GENERAL:

The Elsevier Gender Report (2017) shows that the proportion of women as researchers (2011-2015) in EU-28 reaches 41% for all sciences (aggregated) and 22% for Computer Sciences specifically.

16 The Human Brain Project is one of three EU FET Flagships. “FET Flagships are part of the Future Emerging Technologies (FET) programme, which invests in transformative frontier research and innovation with a high potential impact on technology, to benefit our economy and society.” (see <https://ec.europa.eu/digital-single-market/en/fet-flagships> 25th October 2019)

17 <https://neurobotics.net/partners.html> (5th September 2019)

DOCTORAL GRADUATES:

- Country-level:** For all sciences (aggregated), the share of women doctoral graduates in the mentioned countries varies between 38% (CH) and 52% (IT), while in the fields of studies of ICT the share of women graduates is between 15% (GER) and 25% (IT). (SHE Figures 2018)
- Level of Organisations:** Available data of the organisations show a share of between 13% to 21% of women in computer-related sciences which is a similar range like in the country-benchmark. (SHE Figures 2018)

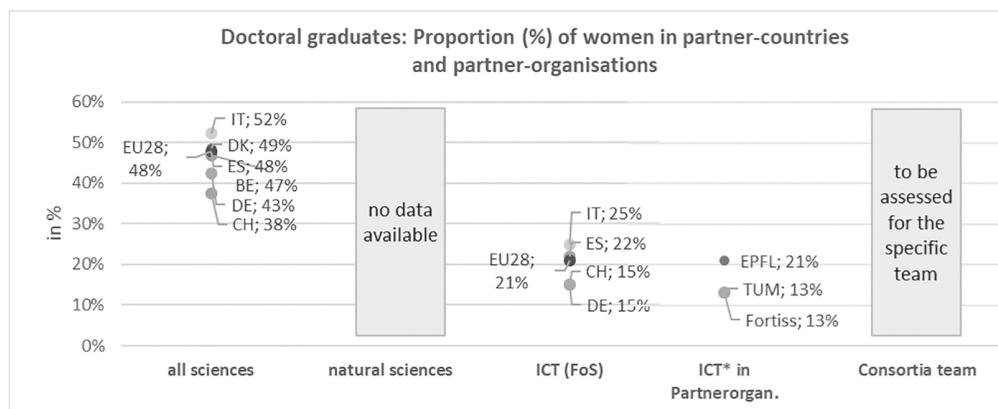


Figure 1: Doctoral graduates: Proportion (%) of women in partner-countries and partner-organisations

Source: She Figures 2018, data for 2016, annual reports of partner-organisations, * TUM/Fortiss: completed promotions in “Informatik”, * EPFL: Computer and Communication Sciences.

PROFESSORS

- Country-level:** The share of women within the career stage of Professors (Grade A¹⁸) varies between 16%-22% for all sciences (aggregated) and is almost the same for natural sciences. No data on narrow fields (e.g. ICT) are published in the “She Figures”. (SHE Figures 2018)
- Level of Organisations:** Data for the organisations show that in ICT-related fields women’s share among professors reached not more than 8% (for TUM C4/W3 Professorships, EPFL full professors).

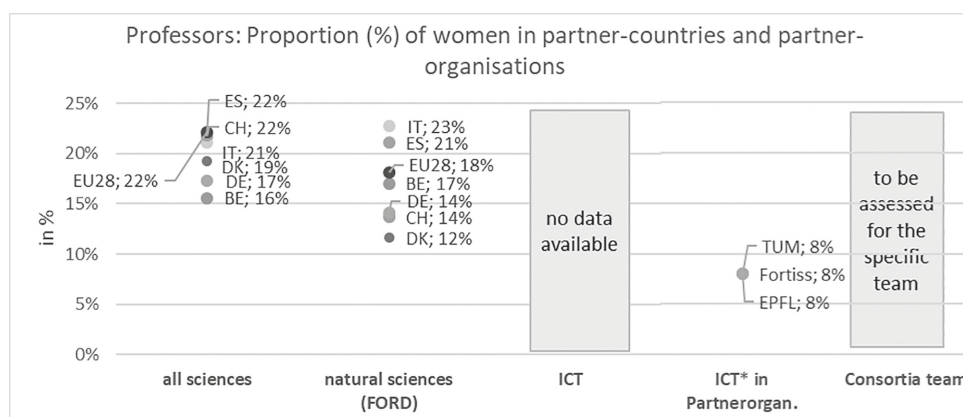


Figure 2: Professors: Proportion (%) of women in partner-countries and partner-organisations

Source: She Figures 2018, data for 2016, annual reports of partner-organisations, * TUM/Fortiss: completed promotions in “Informatik”, *EPFL: PhD-students in “Computer and Communication Sciences”.

PUBLICATIONS (LEIDEN-RANKING)

The Leiden Ranking shows that the number of women authors as a proportion of all publications of the organisations varies for all sciences (aggregated) between 16% and 35% and for “Mathematics and computer science” between 11% and 19%. As mentioned above, this could be used as a rough reference value to compare with the share of women in a specific research or project team.

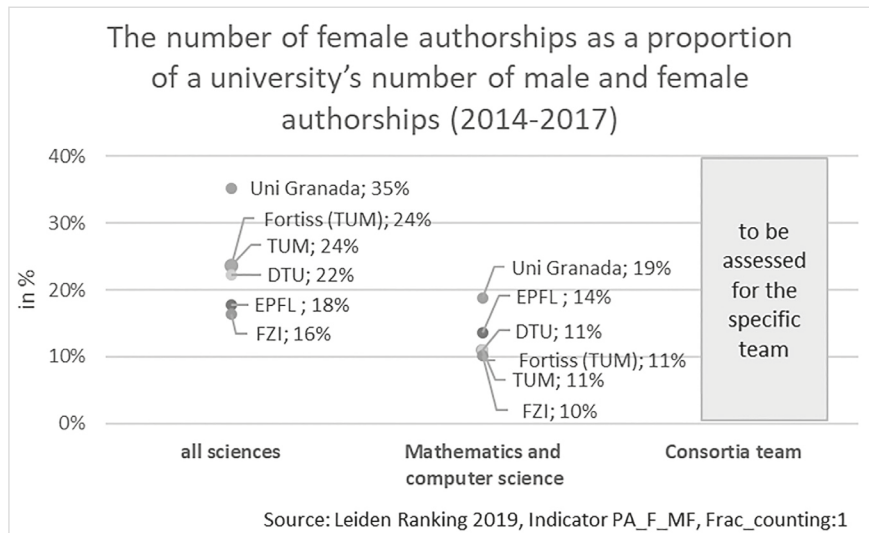


Figure 3: “The number of female authorships as a proportion of a university’s number of male and female authorships (2014-2017)”
Source: Source: Leiden Ranking 2019, Indicator PA_F_MF, Frac_counting:1

CONCLUSIONS OF THE CASE STUDY

It was shown how benchmarks and reference values can be derived for the ICT field and specific research organizations. We summarize our findings on women’s participation as follows:

- While the aggregated proportion of women researchers (2011-2015) in all sciences reaches 41% (EU-28), in the field of Computer Sciences one of five researchers (22%) is a woman.
- For doctoral graduates, data reveal that in ICT fields, women’s proportion is considerably lower than in all sciences aggregated. The partner organisations considered in the case study are no exception. The participation of women is about the same as in the ICT research fields at national level.
- At the professors’ level, the data show that the proportion of women in the natural sciences is slightly lower than in all scientific disciplines (aggregated). For the narrow fields of ICT no national data are available. Among the partner organisations for which data are available, we know that only 8% of ICT professors are women.
- The publication data of the Leiden Ranking show that, depending on the partner organisations, the proportion of women in

publications in the field of “Mathematics and Informatics” is between 10% and 19%. For all partner organisations, the data show that the proportion of women in this scientific field is lower than for all scientific disciplines (aggregated).

Based on these empirical results, the question arises which target values are appropriate to assess the proportion of women contributing in a specific team. The indicators analysed span minimum and maximum values for different career levels (see Table 3 below). To provide the “missing element” and complete the assessment in this case, gender ratios within the actual project team would have to be collected: for the team overall, for leadership positions (professor-level) and doctoral graduates as a basis to compare with the reference values.

A project consortium not reaching minimum values should carefully reflect all possible reasons. It might conclude that some of the partners contribute with specific fields, e.g. related to infrastructure or theoretical computer science in which even less women are actively contributing. If this is not the case, measures to counteract should be taken immediately. In doing so, any support available at the research institutions, or on regional level should be considered.

Career Level	Minimum value		Maximum value	
	% of women	Based on	% of women	Based on
Doctoral and postdoctoral positions	13%	ICT, Partner Institution with lowest value	25%	ICT Graduates in Italy
Principal Investigators	10%	Women as authors in the Leiden Ranking (Mathematics and computer science), partner institution with lowest value	19%	Women as authors in the Leiden Ranking (Mathematics and computer science), partner institution with highest value
Leading position in project consortia or subprojects	8%	Women Professors in ICT at partner organisations	23%	Women professors of natural sciences in Italy

Table 3: Comparative figures for women in ICT at different career levels

3. SUMMARY AND CONCLUSIONS

The aim of this paper is to make an evidence-based contribution to the question of an accurate gender-balance in research teams. We have reviewed existing international data sources that cover gender-specific data in research and assessed the usefulness of these sources. The findings can be summarized as follows:

- A suitable data source for our purpose should fulfill several criteria. It should cover women's proportion among researchers (i) along fields of research, (ii) across countries / organisations or (iii) career levels. Another important question is (iv) if the data source and respective reports are updated regularly.
- There are several sources dealing with gender-related data in research. But as far as we know, none of them covers all mentioned criteria at once. Most sources are severely limited in the availability and comparability of the gender data. This is, among others, due to the complex interplay of different definitions (e.g. regarding fields of study/research or career levels), foci on different levels (e.g. international scientific community in a specific field, national or organizational level) and lacking regularity of reports.
- For our purpose, at least two conditions had to be met in order to be included in a more detailed analysis. The data source had to cover gender-specific data for more than one country as well as for different fields of research or study. According to these criteria, we analysed four sources in detail: (i) the "SHE Figures" report published by the European Commission every three years (latest release "She Figures 2018" in February 2019), (ii) "Gender in the Global Research Landscape" published by Elsevier in 2017, (iii) "UNESCO Science Report" (2015) and (iv) "CWTS Leiden Ranking", which is published annually.
- According to our research, the "She Figures" report has proven to be the most useful database.
- Generally, the more specific the indicators, the less likely it is that current data are actually available. For example: In the "She Figures" report, the shares of women among PhD-students are published at the level of narrow fields of studies for the natural

sciences, ICT and engineering. Data for all other fields of study are only reported in broad categories. With regard to professors, the available data cover only broad fields of research (such as natural sciences, humanities, etc.) and no detailed information is found at all.

- As a consequence, for many research fields, publications like the "She Figures" report will not be enough to derive meaningful data for an assessment of gender-balances in specific research teams, projects or consortia. To overcome this shortage, data must be collected directly from partner organisations to get an idea about women's proportion at different career stages and for specific research fields. Furthermore, an additional literature research focusing on the specific scientific fields of interest can prove useful. Various scientific communities or professional organisations occasionally publish gender-related studies or reports for their specific fields.
- The case study revealed that some, but not all, organisations publish detailed gender-related data. In this paper we limited our research on data which are available online. We suppose that the lack of available data at the level of partner organisations would be reduced within "real" consortia, i.e. when having access to internal data of organisations. Today all organisation monitor their human resources in some way. So, even if the data are not published explicitly, there should at least be internal information within the organisations. Therefore, it should be possible for project partners to collect relevant indicators within their own organisation.

What steps need to be taken if a consortium team wants to assess its own gender relations against benchmarks, as proposed in this paper?

- First, one has to collect the relevant benchmark data. In this paper we have discussed a number of possible data sources as well as existing indicators. We are convinced that a combination of international data sources with bottom-up collected data of research partner organisations can give an appropriate overview of the actual participation of women in specific research fields, as well as on different career levels (PhD, professors). These data can be considered as a benchmark and provide an evidence-based starting point of discussion. They can be com-

pared with a team's or consortium's actual shares of women and men to assess the status quo.

- Second, based on this data-evidence, within a research-team (or an entire organisation), it should be discussed: What is our goal regarding the participation of women and men?
 - Is it enough to be among average countries/organisations in the field? In this case, the average values of the benchmark indicators can be used as **"lower threshold"**.
 - Or do we want to challenge ourselves and set more ambitious goals to reap the potentials of a more balanced team composition? In this case, the research team could follow the best in the field as **"upper threshold"** of the benchmark, or strive to go even beyond. We have referenced the University of Manchester (University of Manchester – Department of Computer Science 2019). Another frequently cited example is the Harvey Mudd College in the USA, where College President Maria Klawe managed to reverse the trend of declining participation of women in computer science (see Harvey Mudd College 2019, Nickelsburg 2019).

This could also be a reasonable requirement by funding agencies, if they adapt the values to the respective scientific field or thematic areas. However, reaching the upper threshold ("best in class") or even an unreflected 50:50 as a binding criterion for research projects does not seem practical. Funding agencies could instead start to compensate those applicants who voluntarily strive for more ambitious goals with additional resources for the necessary additional efforts.

In discussions about an accurate share of women in research teams, it is often argued that the "pool" of suitable researchers is limited. The case study presented in this paper shows for the example of ICT that there is no single and "true" benchmark value. We rather find ranges of values depending on different countries or organisations. For example, the proportion of women PhD researchers in the countries in question varies between 15% and 25% (field of ICT). To estimate if there are really restrictions in the acquisition of women, it is good to be informed about these ranges in the specific fields of research. If the proportion of women in a research team is at the lower end of these ranges, the argument that their share cannot be higher due to a lack of suitable researchers is probably not the whole truth. The persuasiveness of this argument, however, increases when a research team is already positioned at the upper end of the spectrum. Ultimately, it is up to each team to decide how it wants to position itself within an international research community. The prerequisite for this is the knowledge of the actual proportion of women and men in the respective research areas in different countries or research organisations.

REFERENCES / LITERATURE

Bert, Alison (2018). 3 reasons gender diversity is crucial to science. Data reveals that the world needs more women in STEM; here's why. Elsevier Connect, October 1, 2018. <https://www.elsevier.com/connect/3-reasons-gender-diversity-is-crucial-to-science>

Butcher, Gilian; Chiu, Mei-Hung; Gledhill, Igle; Ivie, Rachel; Dawson, Silvina Ponce; Roy, Marie-Francoise (2019). An international and in-

terdisciplinary project to analyze and help reduce the gender gap in mathematics and the natural sciences. AIP Conference Proceedings 2109, 030006 (2019); <https://doi.org/10.1063/1.5110068>

Elsevier (2015). Mapping Gender in the German Research Arena. https://www.elsevier.com/__data/assets/pdf_file/0004/126715/ELS_Germany_Gender_Research-SinglePages.pdf

Elsevier (2017). Gender in the Global Research Landscape. Analysis of research performance through gender lens across 20 years, 12 geographies, and 27 subject areas. Online: https://www.elsevier.com/__data/assets/pdf_file/0008/265661/ElsevierGenderReport_final_for-web.pdf

European Commission (2019). SHE FIGURES 2018. Directorate-General for Research and Innovation. Online: https://ec.europa.eu/info/publications/she-figures-2018_en.

European Union (2019). Gender equality. Website Participant Portal H2020 Online Manual. https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/gender_en.htm

Grasenick, Karin (2019). Counting the numbers: what is a fair gender-balance in science projects? Blogpost from May 24th, 2019. <https://www.ethicsdialogues.eu/2019/05/24/counting-the-numbers-what-is-a-fair-gender-balance-in-science-projects/>

Harvey Mudd College (2019). About Harvey Mudd College – President Maria Klawe. Website. <https://www.hmc.edu/about-hmc/president-klawe/>

Holman, Luke; Stuart-Fox, Devi; Hauser, Cindy E. (2018). The gender gap in science: How long until women are equally represented? PLoS Biol 16(4): e2004956. <https://doi.org/10.1371/journal.pbio.2004956>

Leiden Ranking (2019). CWTS Leiden Ranking 2019. Centre for Science and Technology Studies. Leiden University. Website. <https://www.leidenranking.com/>

National Academy of Sciences, National Academy of Engineering, and Institute of Medicine (2007). Beyond Bias and Barriers: Fulfilling the Potential of Women in Academic Science and Engineering. Washington, DC: The National Academies Press. <https://doi.org/10.17226/11741>.

Netzwerk Frauen- und Geschlechterforschung NRW (2019). Gender-Report der Hochschulen in NRW. Website. <http://www.genderreport-hochschulen.nrw.de/start-genderreport/>

Nickelsburg, Monica (2019). Harvey Mudd President Maria Klawe has a message for computer science grads — and the tech industry. GeekWire Podcast, June 13th, 2019. <https://www.geekwire.com/2019/harvey-mudd-president-maria-klawe-message-computer-science-grads-tech-industry/>

OECD (2015). Frascati Manual 2015. Guidelines for Collecting and Reporting Data on Research and Experimental Development. <https://www.oecd.org/publications/frascati-manual-2015-9789264239012-en.htm>

PLOTINA (2015). Promoting gender balance and inclusion in research, innovation and training. Project funded by EU Horizon 2020 (grant agreement G.A NO 666008). <https://www.plotina.eu/>

Ranga, Marina; Gupta, Namrata; Etzkowitz, Henry (2012). Gender Effects in Research Funding. A review of the scientific discussion on the gender-specific aspects of the evaluation of funding proposals and the awarding of funding http://www.dfg.de/download/pdf/dfg_im_profil/geschaeftsstelle/publikationen/studien/studie_gender_effects.pdf

SAGE (2019). Science in Australia Gender Equity initiative. Website. www.sciencegenderequity.org.au/about-sage/

Schiebinger, Londa; Klinge, Ineke; Sánchez de Madariaga, Inés; Paik, Hee Young; Schraudner, Martina, and Stefanick, Marcia (Eds.) (2011-2018). Gendered Innovations in Science, Health & Medicine, Engineering and Environment. <http://genderedinnovations.stanford.edu/what-is-gendered-innovations.html>

Staats, Cheryl; Capatosto, Kelly; Tenney, Lena & Mamo, Sarah (2017). State of the Science: Implicit bias review. Kirwan Institute. <http://kirwaninstitute.osu.edu/wp-content/uploads/2017/11/2017-SOTS-final-draft-02.pdf>

Technische Universität München (2018). Diversity -Reader der Technischen Universität München. Ausgabe 1/2018. https://www.chancengleichheit.tum.de/fileadmin/w00blt/www/Download/Publikationen/TUM_Diversity_Reader_2018_-_v04.pdf

The Lancet (2019). Advancing women in science, medicine, and global health. Vol. 393, Nr. 10171, p493-610, e6-e28, Feb 09, 2019. [https://www.thelancet.com/journals/lancet/issue/vol393no10171/PIIS0140-6736\(19\)X0006-9#](https://www.thelancet.com/journals/lancet/issue/vol393no10171/PIIS0140-6736(19)X0006-9#)

UNESCO (2015). UNESCO Science Report: towards 2030. https://en.unesco.org/unesco_science_report

UNESCO Institute for Statistics (2019). UIS Statistics Database – Science, technology and innovation. <http://data.uis.unesco.org/>

UNESCO Institute for Statistics (2014). ISCED Fields of Education and Training 2013 (ISCED-F 2013). Manual to accompany the International Standard Classification of Education 2011. <http://dx.doi.org/10.15220/978-92-9189-150-4-en>

University of Manchester – Department of Computer Science (2019). Women in computer science. <https://www.cs.manchester.ac.uk/about/women/>

University of Manchester (UOM) (2019). Equality Information Report 2019. <http://documents.manchester.ac.uk/display.aspx?DocID=35595>

World Economic Forum (2018). The Global Gender Gap Report 2018 (Focus: AI). <https://www.weforum.org/reports/the-global-gender-gap-report-2018>

AUTHORS

MAGDALENA KLEINBERGER-PIERER

convelop cooperative knowledge design gmbh, convelop
Bürgergasse 8-10/I, A-8010 Graz
T: +43 316 720 813
E: Magdalena.Kleinberger-Pierer@convelop.at

SIMON POHN-WEIDINGER

convelop cooperative knowledge design gmbh, convelop
Bürgergasse 8-10/I, A-8010 Graz
T: +43 316 720 813
E: Simon.Pohn-Weidinger@convelop.at

KARIN GRASENICK

convelop cooperative knowledge design gmbh, convelop
Bürgergasse 8-10/I, A-8010 Graz
T: +43 316 720 813
E: Karin.Grasenick@convelop.at

KEYWORDS:

Gender Balance, Higher Education, Indicator, Women, Science

Acknowledgement: Our research has received funding from the European Union Horizon 2020 Framework Programme for Research and Innovation under the Specific Grant Agreement No. 785907 (Human Brain Project SGA2).

EINE UNTERSUCHUNG DER MARKTSITUATION IM BEREICH DER FTI-EVALUIERUNG IN ÖSTERREICH

JÜRGEN STREICHER, WOLFGANG POLT UND MAXIMILIAN UNGER

DOI: 10.22163/fteval.2020.472

INHALTSVERZEICHNIS

1	Einleitung	72
1.1	Hintergrund und Zielsetzung	72
1.2	Aufbau und Methodik	73
2	Entwicklung und Status Quo Der FTI-Evaluierung und -Praxis in	74
2.1	Funktion von Evaluierungen im FTI-Bereich	74
2.2	Rechtliche Rahmenbedingungen bei Vergabeverfahren	74
3	Zentrale Befunde zur Marktsituation im Bereich FTI-Evaluierung in Österreich	75
3.1	Entwicklung und Nachfrage von Evaluierungen	75
3.2	Marktteilnehmer und Marktteilnehmerinnen	77
3.3	Wettbewerbssituation	77
3.4	Vergabepaxis	78
4	Zusammenfassung und Fazit	79
5	Referenzen	80
6	Abbildungsverzeichnis	81
7	Tabellenverzeichnis	81

EINLEITUNG

Diese Studie untersucht die Marktsituation in Bezug auf Evaluierungen im Bereich Forschungs-, Technologie und Innovationspolitik (im Weiteren: FTI-Evaluierungen) in Österreich. Aufbauend auf einer kurzen Darstellung von Entwicklung und Status quo der FTI-Evaluierungskultur in Österreich beleuchtet sie insbesondere die Entwicklung dieses Evaluierungsmarktes, die MarktteilnehmerInnen, die Wettbewerbssituation sowie Vergabepaxis. Die Arbeit schließt mit einer Zusammenfassung der wichtigsten Befunde sowie Schlussfolgerungen für die Praxis der FTI-Evaluierung in Österreich.

1.1 HINTERGRUND UND ZIELSETZUNG

Evaluierungen bzw. evaluierungs-ähnliche Untersuchungen (z.B. Wirkungsanalysen) im Bereich Forschungs-, Technologie und Innovationspolitik (FTI-Politik) sind heute ein häufiger, z.T. auch gesetzlich vorgeschriebener Bestandteil des politischen Prozesses in Österreich. Die Häufigkeit und Qualität von FTI-Evaluierungen hat in den letzten zwei Jahrzehnten deutlich zugenommen. Teilweise wurde in den letzten Jahren aber auch Kritik an der Verfasstheit des österreichischen Evaluierungsmarktes laut. Dieser sei auf zu wenige Anbieter konzentriert, zu wenig international und methodisch zu wenig innovativ. Mangels systematischer Untersuchungen beruhten solche Einschätzungen jedoch hauptsächlich auf subjektiven Einzelwahrnehmungen.

Die Beschaffenheit des österreichischen Marktes für FTI-Evaluierungen war bislang kaum Gegenstand empirisch fundierter Studien. Zwar gibt es Metaevaluierungen zur österreichischen Evaluierungspraxis. Diese fokussieren jedoch auf Art, Methoden, Qualität und Nutzen der untersuchten Evaluierungen (siehe z.B. Landsteiner 2015, Böttcher et al. 2014, Dinges und Schmidmayer 2010). Sie streifen nur am Rande die Beschaffenheit des österreichischen Evaluierungsmarktes, d.h. die Struktur von Auftraggebern und Auftragnehmern (siehe Reiner und Smoliner 2012).

Ziel der vorliegenden Untersuchung ist es, diese „terra incognita“ zumindest teilweise zu erkunden. Im Mittelpunkt stehen Fragen über die Marktgröße bzw. das Marktvolumen, die Marktstruktur sowie der Charakteristika der MarktteilnehmerInnen. Darüber hinaus wurden auch die Einschätzungen der MarktteilnehmerInnen hinsichtlich der Markt- bzw. Wettbewerbssituation sowie relevanter Entwicklungen erhoben.

Die vorliegende Untersuchung soll eine empirische Grundlage für weiterführende Diskussionen schaffen. Naturgemäß konnte sie aufgrund von Zeit- und Ressourcenbeschränkungen nicht alle Aspekte des FTI-Evaluierungsmarktes ausleuchten. So ist etwa der Innovationsgehalt der in Österreich verwendeten Evaluierungsmethoden nicht Gegenstand der vorliegenden Untersuchung gewesen. Ob Marktdynamiken in Österreich anders als in vergleichbaren Ländern verlaufen, konnte mangels Vergleichsdaten ebenfalls nicht festgestellt werden. Es konnten auch nicht alle Aspekte der Wettbewerbsintensität ausgeleuchtet werden. Deshalb wären etwa evaluierungs-ähnliche Untersuchungen, die Erfassung von MarktteilnehmerInnen außerhalb der Plattform fteval oder Vergleiche auf internationaler und regionaler Ebene weitergehende Analysen wert.

1.2 AUFBAU UND METHODIK

Die vorliegende Arbeit untersucht die Marktsituation im Bereich der FTI-Evaluierung in Österreich. Der Fokus liegt also auf dem Bereich Forschung, Technologie und Innovationspolitik (FTI-Politik). Aufgrund der manchmal großen inhaltlichen Nähe zur Hochschulpolitik und zu den damit verbundenen Maßnahmen wurden auch relevante Evaluierungen aus diesem Bereich in die Betrachtungen miteinbezogen.

Bei der Definition des Begriffs „Evaluierung“ wurde auf die Standards der Deutschen Gesellschaft für Evaluation e.V. zurückgegriffen. Unter einer Evaluierung bzw. Evaluation¹ wird demnach die *„systematische Untersuchung von Nutzen und/oder Güte eines Gegenstands (Evaluationsgegenstand) auf Basis von empirisch gewonnenen Daten ... [verstanden. Sie impliziert] ... eine Bewertung anhand offengelegter Kriterien für einen bestimmten Zweck.“* (DeGEval Standards 2016, S. 33). Diese Definition wird auch in den Evaluierungsstandards der Plattform fteval verwendet (fteval Standards 2019, S. 5-6).

Der Untersuchungszeitraum umfasst die zehn Jahre von 2009 bis 2018. Um in die Untersuchung miteinbezogen zu werden, muss bzw. musste der Evaluierungsgegenstand (z.B. ein Programm) einen möglichst klaren Bezug zur FTI-Politik aufweisen – insbesondere im Titel der Untersuchung – sowie auf Bundesebene implementiert sein. Zwar ist es auch möglich, eine Vielzahl von anderen Materien zu evaluieren (Gesetze, Strategien, Richtlinien, Prozesse, Technologien, Dienstleistungen, Veranstaltungen, Schulungen, Materialien, Berichte u.a., vgl. fteval Standards 2019, S. 12) – die entsprechende Berücksichtigung dieser Gegenstände hätte jedoch den Rahmen der gegenständlichen Arbeit gesprengt.

Kapitel 2 bietet einen kurzen Überblick über die Entwicklungen und den Status quo der FTI-Evaluierungskultur und -praxis in Österreich. Dabei werden Funktionen sowie die rechtlichen Rahmenbedingungen von Evaluierungen im FTI-Bereich sowie die rechtlichen Rahmenbedingungen beschrieben. Kernstück der Studie bilden die empirischen Befunde, welche auf Basis von Sekundärdaten (Überblick vgl. Tab. 1) sowie einer Befragung von MarktteilnehmerInnen gewonnen wurden (siehe Kapitel 3).

2009-2018	Repository der Plattform für Forschungs- und Technologiepolitikevaluierung (fteval) Veröffentlichungen der Bundesforschungsdatenbank (bfdat) Internetauftritte der relevanten Ministerien und Förderagenturen sowie öffentliche Mittler (auftrag.at, Bundesbeschaffung GmbH) und Evaluierungs-Anbieter in Österreich (fteval Mitglieder)
2014-2018	Beantwortung parlamentarischer Anfragen betreffend die Erbringung von Dienstleistungen für das BMBWF, BMVIT und BMDW bzw. relevante Vorgängereinrichtungen Vergabebekanntmachungen im „Supplement zum Amtsblatt der Europäischen Union (Tenders Electronic Daily, TED)“

Tabelle 1: Verwendete Quellen von Sekundärdaten

Auf Basis der Sekundärdaten wurde die Anzahl von Evaluierungen im FTI-Bereich von 2009 bis 2018 recherchiert. Evaluierungen, die in mehreren Schritten durchgeführt wurden bzw. mehrere zusammenhängende Berichte umfassten, wurden dabei als eine Untersuchung gewertet. Die verschiedenen Informationen zu den einzelnen Evaluierungen, z.B. Titel, Abschlussjahr, AuftragnehmerInnen etc. wurden in einer Liste zusammengeführt und den identifizierten AuftraggeberInnen mit der Bitte um Vervollständigung zur Verfügung gestellt.

Im Zeitraum Mai bis Juni 2019 wurden die Mitgliedseinrichtungen der Plattform fteval gebeten, an einer Online-Befragung, die sich jeweils an AnbieterInnen bzw. AuftraggeberInnen von Evaluierungen richtete, teilzunehmen. Bei Anbietern lag der Schwerpunkt dieser Erhebung auf dem Umfang ihrer Tätigkeiten im Evaluierungsbereich sowie auf dem Stellenwert dieses Bereichs in ihrem Leistungsportfolio. Bei Nachfragern lag der Schwerpunkt auf der Anzahl eingelangter Bewerbungen und vergebener Aufträge sowie auf den vergebenen Volumina. Die Fragen bezogen sich überwiegend auf die letzten drei Jahre (2016 bis 2018). Die erhobenen Inhalte wurden mit den Ergebnissen der Sekundärdatenanalyse abgeglichen sowie mit den Einschätzungen von Mitgliedern der Plattform fteval im Rahmen einer Diskussionsveranstaltung² kontrastiert.

Als für die Befragung relevante MarktteilnehmerInnen wurden 13 potentielle Auftragnehmer und zehn potentielle AuftraggeberInnen identifiziert. Diesen wurden Links zur Online-Befragung zugesandt. Die Versendung erfolgte nicht durch die JOANNEUM RESEARCH sondern durch eine Rechtsanwaltskanzlei. Dies sicherte die Anonymität der Angaben und gewährleistete, dass die Studienautoren keinen Einblick in die Situation einzelner MitbewerberInnen erhielten. Von den so kontaktierten Marktteilnehmern übermittelten 12 Auftragnehmer und acht AuftraggeberInnen einen auswertbaren Fragebogen. Der Rücklauf lag somit bei rund 90 % bzw. 80 %.³ Die Studienautoren möchten an dieser Stelle den Mitgliedern der Plattform fteval nochmals herzlich für die Teilnahme an der Befragung danken.

1 Die Begriffe Evaluierung und Evaluation wurden im Rahmen der Sekundärdatenanalyse und Primärerhebung synonym verwendet. Die vorliegende Arbeit wählt Evaluierung als primäre Bezeichnung.
2 „Der Evaluierungsmarkt im FTI-Bereich in Österreich. Ergebnisse der Marktstudie im Auftrag der Österreichischen Plattform für Forschungs- und Technologiepolitikevaluierung“. Veranstaltung der Plattform fteval, 5. Juli 2019, Wien.
3 Rücklauf AuftraggeberInnen: Neun, auswertbar waren acht: Aufgrund der geringen Überlappung von durchführenden bzw. beauftragenden Einrichtungen lässt sich abschließend nicht klären, ob nicht doch eine Vollerhebung stattgefunden hat.

Bezüglich der Analyse der Befragungsergebnisse ist zu beachten, dass aufgrund der Zusammensetzung der Mitgliedseinrichtungen der Plattform fteval zwar eine repräsentative, aber kleine Samplegröße⁴ zur Verfügung stand, sodass deskriptive Auswertungen gewissen Verzerrungen durch Ausreißer unterliegen können.

In einem abschließenden Arbeitsschritt wurden die Ergebnisse zusammengeführt und einer analytischen Betrachtung unterzogen (siehe Kapitel 4).

2. ENTWICKLUNG UND STATUS QUO DER FTI-EVALUIERUNGSKULTUR UND -PRAXIS IN ÖSTERREICH

Evaluierungen haben sich als Instrument im Prozess der FTI-Politik international etabliert. Sie ermöglichen strukturiertes Feedback und die Bereitstellung politikrelevanten Wissens, nicht nur für die Planung, Entwicklung und Bewertung bestehender Maßnahmen. Sie schaffen auch verbesserte Grundlagen für die Entwicklung von Ideen und die Erarbeitung neuer Initiativen. Nicht zuletzt im Kontext der europäischen FTI-Politik steigt der Informations- und Lernbedarf über die Funktionsweise des österreichischen Forschungs- und Innovationssystems, die Qualität der österreichischen Forschung und Entwicklung und die Wirksamkeit von FTI-Förderung im internationalen Vergleich.

2.1 FUNKTION VON EVALUIERUNGEN IM FTI-BEREICH

Vor diesem Hintergrund sind in den letzten zwei Jahrzehnten die FTI-Evaluierungstätigkeiten in Österreich stark angestiegen. Evaluierungskompetenzen und -kapazitäten im öffentlichen Bereich wurden ausgebaut. Entsprechende Zuständigkeiten in Ressorts und Agenturen wurden geschaffen. Österreich liegt heute hinsichtlich der Anzahl von Evaluierungen im FTI-Bereich im europäischen Spitzenfeld. Studien, die sich dem Umgang mit Evaluierungen widmen, unterstreichen die im Allgemeinen hohe Professionalität und Qualität österreichischer Evaluierungen (vgl. Tsiouri und Sidiropoulos 2014, Dinges und Schmidmayer 2010; siehe auch Reiner und Smoliner 2012, BMWFW und BMVIT 2017). Andererseits ist eine zunehmende Institutionalisierung und Routinisierung zu beobachten, mit Auswirkungen auf das Nutzen von und das Lernen aus Evaluierungen (vgl. Streicher 2017, Landsteiner 2015, Biegelbauer 2013).

Die österreichische Evaluierungspraxis im FTI-Bereich spiegelt die Struktur des Forschungsförderungssystems wider. Diese Evaluierungs-

praxis ist von Evaluierungen zahlreicher, relativ kleiner einzelner Programme und Maßnahmen geprägt. Institutionen, Organisationen oder gesetzliche Regelungen werden dagegen selten evaluiert. Evaluierungen nehmen überwiegend eine formative Perspektive ein und finden in bestimmten, oft gesetzlich vorgeschriebenen, Intervallen während der Laufzeit einer Intervention statt (sogenannte Interims- oder Zwischenevaluierungen). Ex-ante Evaluierungen, begleitende Evaluierungen bei erstmaliger Umsetzung und summative Evaluierungen sind weniger häufig anzutreffen.

Evaluierungen fungieren als wichtiges Instrument der FTI-Politik und Verwaltungsführung, das dazu beiträgt, Transparenz, Rechenschaftslegung und eine evidenzbasierte Entscheidungsfindung zu unterstützen. Die Durchführung erfolgt in Österreich auf der Basis allgemeiner gesetzlicher Erfordernisse, spezifischer Anforderungen im Kontext von Richtlinien und Fördertätigkeiten sowie haushaltrechtlicher Maßgaben. Die Rechtsgrundlagen wurden in den letzten Jahren laufend aktualisiert, Begrifflichkeiten angepasst sowie Regelungen überarbeitet. Zu den maßgeblichen Gesetzen und Vorschriften zählen das Forschungs- und Technologieförderungsgesetz (FTF-G), die Allgemeinen Rahmenrichtlinien für die Gewährung von Förderungen aus Bundesmitteln (ARR 2014), das Forschungsorganisationsgesetz (FOG; Berichtswesen: §§ 6–9) sowie die auf diesen Gesetzen basierenden Richtlinien zur Forschungsförderung und zur Förderung⁵ der wirtschaftlich-technischen Forschung, Technologieentwicklung und Innovation (die sogenannten FTI-Richtlinien).⁶ Im Einklang mit der Implementierung der wirkungsorientierten Haushalts- und Verwaltungsführung wurden 2015 die Richtlinien überarbeitet. Dabei wurde eine Fokussierung auf inhaltliche Ziele und Indikatoren vorgenommen (vgl. BMBWF, BMVIT und BMDW 2018).

Zudem entwickelte sich in den letzten Jahren ein reger Diskurs über Möglichkeiten, Funktion und Nutzen von Evaluierungen, Ansprüche an und den Umgang mit Evaluierungen (vgl. BMWFW und BMVIT 2017, siehe auch: Warta und Philipp 2016). Es hat sich rund um die Evaluierungspraxis eine Community mit einer „gemeinsamen Sprache“ entwickelt. Dort wird über Themen wie Wirkungsorientierung und Folgenabschätzung debattiert. Die Plattform fteval stellt dabei seit Jahren ein wichtiges Forum dar. Sie formuliert und publiziert Standards der Evaluierung in der Forschungs- und Technologiepolitik (vgl. fteval Standards 2019).

2.2 RECHTLICHE RAHMENBEDINGUNGEN BEI VERGABEVERFAHREN

Evaluierungen stellen aus der Sicht der öffentlichen Hand einen Beschaffungsfall dar. Sie werden als Bestellungen der öffentlichen Hand im Rahmen der öffentlichen Auftragsvergabe realisiert. In Österreich können öffentliche Aufträge durch Bedarfsbündelung der öffentlichen

4 Das betrifft insbesondere die Angaben der AuftraggeberInnen von Evaluierungen, die sich in der Plattform fteval aus verschiedenen Einrichtungen mit unterschiedlichem Nachfrageverhalten nach Evaluierungen zusammensetzen.

5 Vgl. Richtlinien der Bundesregierung über die Gewährung und Durchführung von Förderungen gemäß §§ 10–12 FOG, BGBl. Nr. 341/1981

6 Vgl. Richtlinien zur Förderung der wirtschaftlich-technischen Forschung, Technologieentwicklung und Innovation (FTI-Richtlinie 2015); das sind: Themen-FTI-Richtlinie, Struktur-FTI-Richtlinie, Humanressourcen-FTI-Richtlinie gemäß Forschungs- und Technologieförderungsgesetz (FTF-G) des Bundesministers für Verkehr, Innovation und Technologie (GZ BMVIT-609.986/0011-III/12/2014) und des Bundesministers für Wirtschaft und Arbeit (GZ BMWFW-97.005/0003-C1/9/2014).

Einrichtungen zentral oder durch jeden Bedarfsträger selbst, also dezentral, im Rahmen der Gesetze, vergeben werden.⁷ Die zentrale Beschaffung für Bundesdienststellen und andere öffentliche AuftraggeberInnen übernimmt in Österreich seit 2001 die Bundesbeschaffung GmbH (BBG) als Tochter des Bundesministeriums für Finanzen (BMF).

Seit dem Beitritt zur Europäischen Union (EU) 1995 existiert in Österreich für die öffentliche Auftragsvergabe ein gesetzliches Regulativ. Seit 2002 gilt ein einheitliches Bundesvergabegesetz (BVerG). Zuletzt wurde dieses Gesetz 2018 novelliert. Zu den zentralen Grundsätzen des Vergaberechts (§ 20 BVerG 2018) zählen u.a. ein freier und lauterer Wettbewerb, Gleichbehandlung aller Bewerber sowie allgemein die Gewährleistung von Transparenz in Vergabeverfahren.

Das Bundesvergabegesetz sieht in seiner aktuellen Fassung eine Reihe von Vergabeverfahren vor (§ 31 BVerG 2018). Diese sind: offenes Verfahren, nicht offenes Verfahren mit/ohne vorherige(r) Bekanntmachung, Verhandlungsverfahren mit/ohne vorherige(r) Bekanntmachung, Direktvergabe (auch: mit vorheriger Bekanntmachung), Rahmenvereinbarung, dynamisches Beschaffungssystem, wettbewerblicher Dialog, elektronische Auktion (kein eigenes Vergabeverfahren) und die Innovationspartnerschaft. Weiters lassen sich die Verfahrensarten nach den durchzuführenden Schritten (ein- bis dreistufiges Vergabeverfahren) unterteilen.

Wesentliches Kriterium für die Art der Vergabe ist der geschätzte Auftragswert der zu vergebenden Leistung. Im Vergabeverfahren wird zwischen dem Oberschwellenbereich und dem Unterschwellenbereich unterschieden. Im Oberschwellenbereich kommen nicht sämtliche Verfahrensarten zur Anwendung, und es bestehen auch höhere Anforderungen an die Publizität des Vergabeverfahrens. Ob ein Vergabeverfahren im Oberschwellenbereich oder Unterschwellenbereich vorliegt, orientiert sich am (geschätzten) Auftragswert. Der Grenzwert zum Oberschwellenbereich für öffentliche Liefer- und Dienstleistungsaufträge einer zentralen Beschaffungsstelle sowie vergleichbarer Bundeseinrichtungen beträgt 144.000 € (§ 12 BVerG 2018). Für alle übrigen Liefer- und Dienstleistungsaufträge beläuft sich der Schwellenwert auf 221.000 € (Baufträge: 5.548.000 €). Ab dem Erreichen dieser Grenzwerte muss zusätzlich zur Bekanntmachung im regionalen Amtsblatt eine Meldung an das Amtsblatt der europäischen Union (Tenders Electronic Daily, TED) ergehen.

Im Unterschwellenbereich wurde die Schwellenwertverordnung von 2009 (VO BGBl. II Nr. 125/2009) im Jahr 2018 bis zum 31. Dezember 2020 neuerlich verlängert. Aufträge im Bau-, Liefer- und Dienstleistungsbereich können weiterhin bis zu einem geschätzten Auftragswert von 100.000 € (ohne Ust.) direkt, ggf. nach Einholung von Angeboten oder unverbindlichen Preisauskünften, von einem oder mehreren Interessenten an geeignete leistungsfähige und zuverlässige AuftraggeberInnen vergeben werden. Anzumerken ist, dass in den relevanten Ressorts verschiedene, mehr oder weniger formalisierte, Verfahren und Vorgangsweisen etabliert wurden, welche die Direktvergabe von geistigen Dienstleistungen unter Berücksichtigung der Grundsätze des Vergaberechts regeln. Dazu gehören u.a. Bestimmungen bezüglich der Einholung und Bewertung von Angeboten.

3. ZENTRALE BEFUNDE ZUR MARKTSITUATION IM BEREICH FTI-EVALUIERUNG IN ÖSTERREICH

Im Folgenden werden die wichtigsten Befunde aus der Analyse der Sekundärdaten und der Online-Befragung präsentiert.

3.1 ENTWICKLUNG UND NACHFRAGE VON EVALUIERUNGEN

Im Zuge der ersten Recherchephase konnten für den Zeitraum 2009 bis 2018 über 500 Studien, Analysen und Expertisen mit mehr oder weniger starkem Bezug zum Themenfeld FTI identifiziert werden. Unter Berücksichtigung der eingangs beschriebenen Definitionen konnte die Anzahl auf 107 einzelne Evaluierungen eingegrenzt werden. Abb. 1 gibt die Entwicklung der Zahl dieser Evaluierungen im Zeitverlauf wider. Diese zeigt starke Schwankungen bis 2014, was zum einen auf die umfassende Systemevaluierung zu Beginn des Betrachtungszeitraums (Aiginger et al. 2009) und die im Vergleich zu den Folgejahren schwächere Datenbasis zurückgeführt werden kann. Andererseits ist zu vermuten, dass die Sekundärdatenbasis für frühere Jahre schwächer ist, da ein Trend zur vermehrten Darstellung von Referenzprojekten auf den Websites der Anbieter feststellbar ist.

Über den gesamten Zeitraum betrachtet lag die Leitung der Evaluierungen in ca. neun von zehn Fällen bei einer österreichischen Einrichtung. Der Rest verteilt sich auf Einrichtungen aus anderen europäischen Ländern, insbesondere auf solche aus Deutschland. Knapp drei Viertel aller Evaluierungen wurden von Einrichtungen durchgeführt, die mit Stand Juni 2019 auch Mitglied bei der Plattform fteval waren. In einem Drittel der übrigen Fälle waren diese, soweit dies nachvollziehbar war, immerhin als Partner dabei. Der Anteil an fteval-Mitgliedern an den durchgeführten Evaluierungen ist in den letzten Jahren des Beobachtungszeitraums jedoch zurückgegangen.

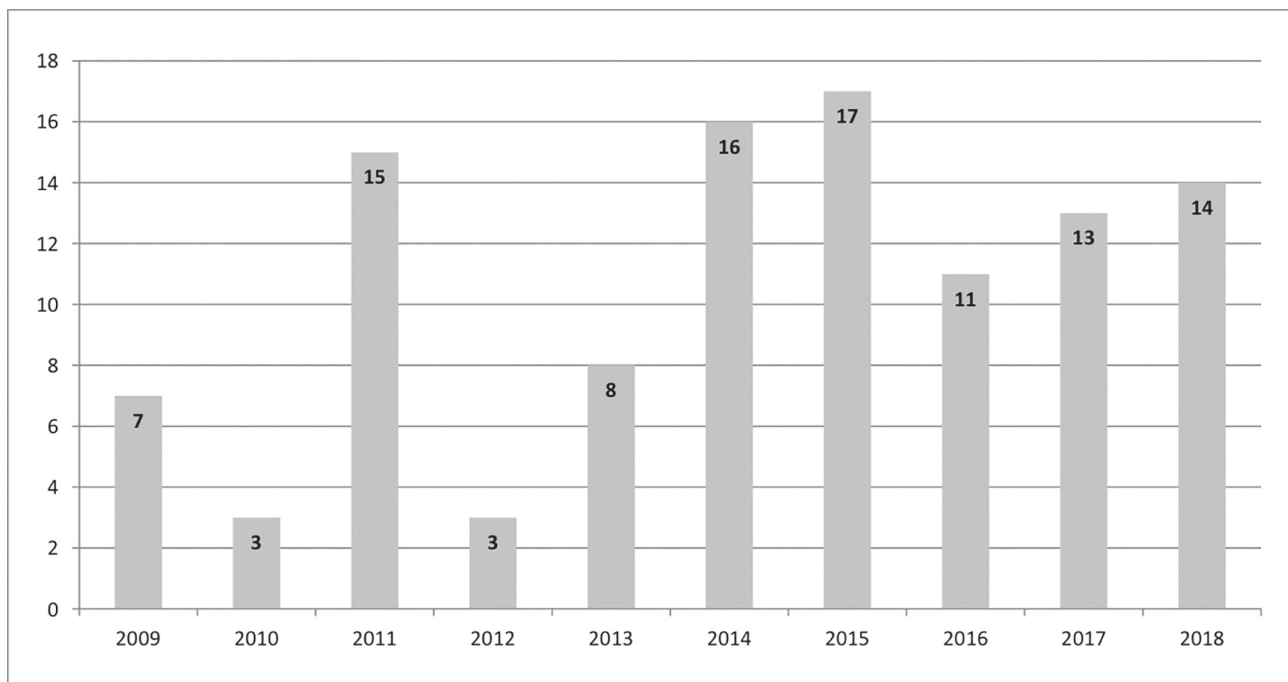


Abbildung 1: Anzahl an identifizierten Evaluierungen im Zeitraum 2009-2018

Quelle: Sekundärdatenanalyse (vgl. Tab. 1).

Ein Vergleich mit einer der frühesten Erhebungen (Dinges und Schmidmayer 2010) zeigt, dass sich die Evaluierungsaktivitäten seit Mitte der 2000er Jahre ca. verdoppelt haben: Wurden zwischen 2002 und 2007 im Schnitt sechs Evaluierungen pro Jahr durchgeführt, so waren es laut Sekundärdatenanalyse im Zeitraum 2016 bis 2018 bereits 13 (vgl. Tab 2).

Laut den AuftraggeberInnen wurden im Untersuchungszeitraum im Schnitt rund 17 Evaluierungen pro Jahr beauftrag, was insgesamt als Untergrenze angenommen werden kann. Die Differenz zum Ergebnis der Sekundärdatenanalyse ist zum Teil auf nicht publizierte Evaluierungen zurückzuführen. Die höchste Anzahl an durchgeführten Evaluierungen (25) im Zeitraum 2016 bis 2018 wird in der Befragung der AuftragnehmerInnen genannt. Es ist zu vermuten, dass hier auch Arbeiten berücksichtigt wurden, die nicht nur Maßnahmen auf Bundesebene adressieren oder

eher dem Spektrum evaluierungs-ähnlicher Untersuchungen zuzuordnen sind – aber durch die AuftragnehmerInnen als Evaluierungen im Sinne dieser Untersuchung eingestuft wurden.

Mit Blick auf die unterschiedlichen Arten von Evaluierungen (siehe Tab. 2) fällt erwartungsgemäß die Dominanz von Interims- bzw. Zwischenevaluierungen auf. Überraschend ist die von den AuftragnehmerInnen genannte hohe Anzahl an Ex-ante Evaluierungen. Dieser Befund steht im Kontrast zu bisherigen Beobachtungen. Abseits der oben angeführten Vermutung bezüglich des Umfangs der berücksichtigten Arbeiten dürften verwandte und dennoch unterschiedliche Konzepte (z.B. Roadmaps, Entwicklungsszenarien, Foresight und Technikfolgen-Abschätzungen), die zuletzt verstärkt Einsatz im FTI-Bereich fanden (vgl. BMWFW und BMVIT 2017) – in der Befragung als Ex-ante Evaluierungen eingeordnet worden sein.

Evaluierungsart ¹	Durchschnitt pro Jahr			
	2002-2007 ^a	2016-2018 ^b	2016-2018 ^c	2016-2018 ^d
Ex-ante Evaluierungen ²	0,8	0,3	0,7	4,3
Interims-/Zwischenevaluierungen ³	0,7	10,7	7,3	12,0
Begleitende Evaluierungen ⁴	3,0	1,0	3,3	4,3
Ex-post Evaluierungen (inkl. andere) ⁵	1,2	0,7	5,3	4,3
Summe	5,7	12,7	16,7	25,0

Tabelle 2: Evaluierungen auf nationaler Ebene, 2002-2007 sowie 2016-2018

Anmerkungen: ^a Dinges und Schmidmayer 2010, ^b Sekundärdatenanalyse, ^c Befragung AuftraggeberInnen, ^d Befragung AuftragnehmerInnen; ¹ Definitionen laut fteval Standards 2019, S. 12; vergleichbar mit den Definitionen in Dinges und Schmidmayer 2010, S. 253. ² setzen vor Beginn einer Intervention an und haben einen zukunftsgerichteten Charakter; ³ durchgeführt während der Laufzeit einer Intervention, unterschiedliche Foki; ⁴ setzen zu Beginn einer Intervention an und erfolgen parallel zu Ihrer Umsetzung, unterschiedliche Foki; ⁵ direkt und/oder lange nach Abschluss einer Intervention, inkl. „terminale“ Evaluierungen zum Maßnahmenende bzw. „other“ laut Dinges und Schmidmayer 2010.

Summiert über alle befragten AuftragnehmerInnen, geben diese an, zwischen 2016 und 2018 im Schnitt pro Jahr ca. 750.000 € durch FTI-Evaluierungen eingeworben zu haben. AuftraggeberInnen schätzen das Volumen an extern vergebene Evaluierungen für denselben Zeitraum im Jahresdurchschnitt auf ca. 760.000 €. Die Einschätzungen von AuftraggeberInnen und AuftragnehmerInnen zur Marktreife decken sich also in hohem Maße. Die Ergebnisse der Sekundärdatenanalyse sind diesbezüglich unvollständig, weisen aber tendenziell in dieselbe Richtung.

Im Einklang mit den Befunden der Sekundärdatenanalyse geben 58 % (sieben von zwölf) der befragten AuftragnehmerInnen an, dass die Nachfrage gemessen am Gesamtvolumen der ausgeschriebenen Evaluierungsaufträge im FTI-Bereich im Inland zwischen 2016 und 2018 unverändert geblieben ist (für 25 % zugenommen, für 17 % abgenommen). Laut 63 % (fünf von acht) der befragten AuftraggeberInnen war im selben Zeitraum in ihren Einrichtungen die Nachfrage nach Evaluierungen unverändert (für 37 % zugenommen). Zuletzt entwickelte sich die Nachfrage nach Evaluierungen eher gleichmäßig, was sich mit jüngsten Beobachtungen über die Regelmäßigkeit und Routine des Evaluierungsaufkommens im FTI-Bereich in Österreich deckt (Streicher 2017, Landsteiner 2015).

3.2 MARKTTEILNEHMER UND MARKTTEILNEHMERINNEN

Institutionell wird das Politikfeld FTI vor allem durch drei Ministerien bestimmt: (1) Bundesministerium für Bildung, Wissenschaft und Forschung (BMBWF), (2) Bundesministerium für Digitalisierung und Wirtschaftsstandort (BMDW) und (3) Bundesministerium für Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie (BMK). Diese bzw. relevante Vorgängereinrichtungen sind, oft gemeinsam auftretend, aktuell die wesentlichen AuftraggeberInnen für Evaluierungen auf Bundesebene. Die Sekundärdatenanalyse bestätigt dies. Weitere relevante Ressorts für die Vergabe von FTI-Evaluierungen, die jedoch nicht befragt wurden, können das Bundesministerium für Finanzen (BMF) und das Bundesministerium für Soziales, Gesundheit, Pflege und Konsumentenschutz (BMSGPK) sein. Auch das Bundesministerium für Landwirtschaft, Regionen und Tourismus (BMLRT) und das Bundesministerium für Landesverteidigung (BMLV) beauftragen gelegentlich Evaluierungen bzw. evaluierungs-ähnliche Untersuchungen mit Forschungs- und/oder Technologieinhalten.

Die Gegenstände von Evaluierungen – oft FTI-Programme – werden zumeist von Agenturen im Auftrag eines Trägers/Ministeriums umgesetzt. Diese Agenturen sind im Bereich der angewandten Forschung die Forschungsförderungsgesellschaft (FFG) sowie das Austrian Wirtschaftsservice (aws). Im Bereich der Grundlagenforschung ist es der Fonds zur Förderung der wissenschaftlichen Forschung (FWF). Laut Sekundärdatenanalyse tritt der FWF für Evaluierungen, die seine Programme und Maßnahmen betreffen, regelmäßig auch als Auftraggeber auf. Im Gegensatz dazu ließen sich nur wenige Evaluierungen in den Sekundärdatenquellen finden, die von FFG oder aws beauftragt wurden.

Die befragten AnbieterInnen von Evaluierungen unter den fteval-Mitgliedsorganisationen, welche auch am Großteil der Evaluierungen der letzten Jahre beteiligt waren, setzen sich aus unterschiedlichen Forschungsorganisationen und Beratungsunternehmen zusammen (vgl. Warta und Philipp 2016). Laut Befragung stuft sich die Hälfte (sechs von zwölf) als private außeruniversitäre Forschungseinrichtung ein, der Rest ordnet sich öffentlichen außeruniversitären Forschungseinrichtungen,

privatwirtschaftlichen Unternehmen sowie sonstigen Einrichtungen zu. Gemessen am Gesamtumsatz zeigt sich eine relativ gleichmäßige Verteilung der Einrichtungen über ein Spektrum von „kleinen“ mit weniger als 500.000 € bis hin zu „großen“ Einrichtungen mit mehr als 5 Mio. € Gesamtumsatz pro Jahr, es sind also Einrichtungen unterschiedlichster Größe auf dem österreichischen Markt vertreten.

Die AuftragnehmerInnen sind vielfach international aktiv: Zwei Drittel (acht von zwölf) geben an, in ausgewählten europäischen Ländern tätig zu sein. Für mehr als die Hälfte stellt die gesamte EU einen Markt dar. Auch die Themenbereiche, in welchen zwischen 2016 und 2018 Evaluierungen angeboten bzw. durchgeführt wurden, sind breit gestreut. Kaum abgedeckt werden durch die hier erfassten Institutionen aber Themen der internationalen Entwicklung oder der Kultur und Kulturpolitik.

Zu den am häufigsten genannten evaluativen Tätigkeitsschwerpunkten der befragten Einrichtungen zählen die Evaluierung von Förder- und Forschungsprogrammen sowie die Erstellung von Studien, die in Teilen einen Evaluierungscharakter aufweisen (jeweils 80 %, zehn von zwölf). Bei der Evaluierung von Einrichtungen bzw. Institutionen beträgt der Anteil 75 % (neun von zwölf), bei der Durchführung von Wirkungsanalysen 67 % (neun von zwölf). Reviews und Gutachten werden relativ selten als Tätigkeitsschwerpunkt genannt.

Nur wenige Einrichtungen bilden das gesamte Spektrum an evaluativen Tätigkeiten ab. Viele haben ein über die Jahre entwickeltes Set an Tätigkeitsschwerpunkten und können auf Erfahrung in mehreren Themenbereichen aufbauen. Nahezu alle befragten AuftragnehmerInnen sind zudem in den letzten Jahren im Zuge von FTI-Evaluierungen eine Kooperation mit anderen Partnern eingegangen.

Für die befragten Auftragnehmer spielen FTI-Evaluierungen im Rahmen ihrer Geschäftstätigkeit eine wichtige, jedoch nicht die Hauptrolle. 2018 lag der Anteil an entsprechenden Evaluierungen am Gesamtbudget bzw. -umsatz bei 27 % (im Inland vergeben) bzw. 15 % (im Ausland vergeben). Für die nächsten Jahre (bis 2023) wird eine leichte Verschiebung der Anteile zugunsten von Evaluierungen, die im Ausland vergeben werden, erwartet. Die erwirtschafteten Erträge aus Evaluierungsaufträgen werden als überwiegend kostendeckend eingeschätzt (75 %, neun von zwölf). Ein Viertel gab jedoch auch an, dass Evaluierungen, die auf nationaler Ebene vergeben wurden, in Summe nicht kostendeckend durchgeführt werden konnten.

3.3 WETTBEWERBSSITUATION

Die Ausschreibung und Vergabe von Aufträgen im Bereich der FTI-Evaluierungen in Österreich sehen 67 % (acht von zwölf) der befragten AuftragnehmerInnen als stark auf wenige AuftraggeberInnen konzentriert. Dabei verfügen lediglich 25 % (drei von zwölf) über Schlüsselauftraggeber bzw. Stammkunden. Bei 42 % (fünf von zwölf) wechseln die AuftraggeberInnen häufig, bei 33 % (vier von zwölf) zumindest gelegentlich. Etwas anders stellt sich die Sichtweise bei den AuftraggeberInnen dar: 62 % (fünf von acht) sehen Ihre Einrichtung als eine unter vielen, lediglich 25 % (zwei von acht) als eine unter den wenigen wichtigen Nachfragern. Übereinstimmend ist das Bild dagegen hinsichtlich der Zusammensetzung der AuftragnehmerInnen im Laufe der Jahre: Diese wechselten auch laut Einschätzung der AuftraggeberInnen zuletzt häufig (für 50 %) bzw. gelegentlich (für 37 %).

Während 25 % der befragten AuftragnehmerInnen den Wettbewerb auf nationaler Ebene als zuletzt gleichbleibend einschätzen, attestieren

67 % (acht von zwölf) eine verschärfte Wettbewerbssituation, wobei diese auf internationaler Ebene noch ausgeprägter wahrgenommen wird (82 %). Auch 88 % (sieben von acht) der befragten AuftraggeberInnen schätzen den Wettbewerb im Inland als sehr stark ein (vgl. Abb. 2).

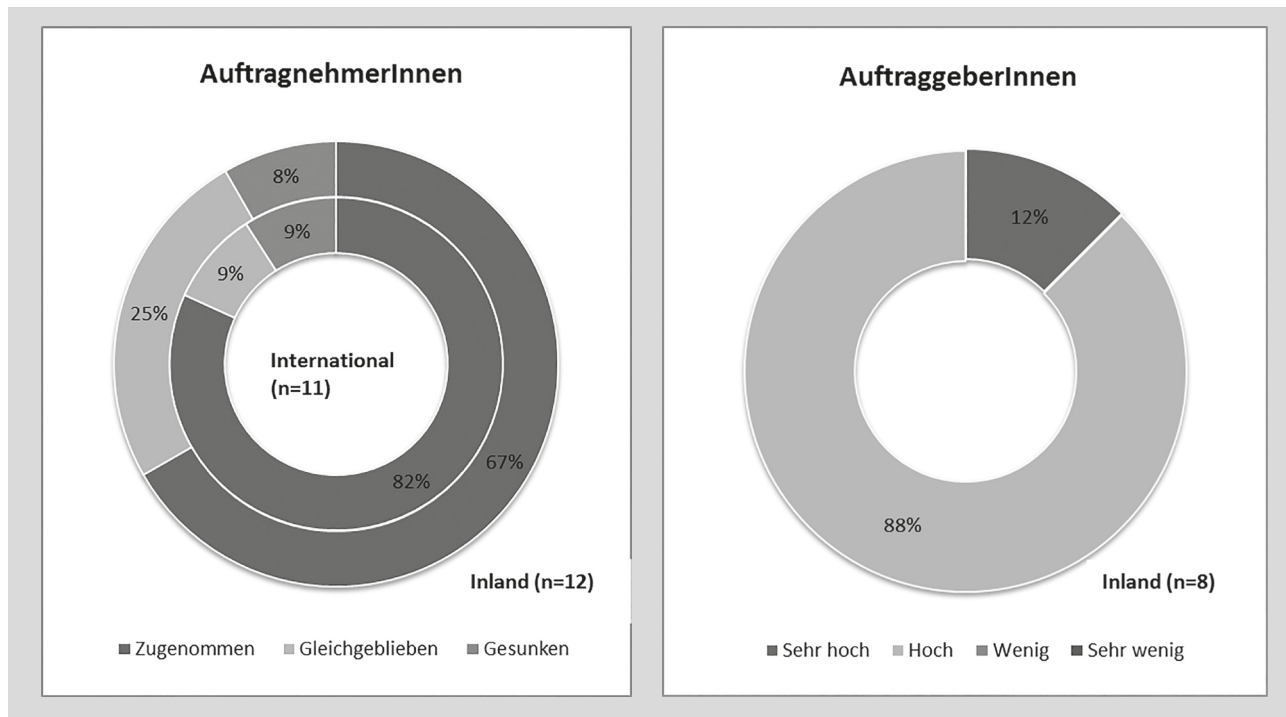


Abbildung 2: Einschätzung der Wettbewerbssituation, 2016-2018

Quelle: Online-Befragung AuftragnehmerInnen und AuftraggeberInnen 2019. Werte sind arithmetische Mittel.

Als Gründe für die verschärfte Wettbewerbssituation im Inland werden von den befragten Auftragnehmern am häufigsten der Markteintritt neuer MitbewerberInnen sowie ein stärkerer Preiswettbewerb genannt. Mit Blick auf die Entwicklungen der letzten fünf bis zehn Jahre wurden zudem steigende Anforderungen bei gleichbleibenden Budgets angeführt. Budgetäre Einschränkungen gehen laut Wahrnehmung der AuftragnehmerInnen jedoch auch mit dem Verzicht auf weitreichendere, und somit kostenintensivere Fragestellungen in Leistungsbeschreibungen einher.

Das gezeichnete Bild deckt sich zum Großteil mit den Einschätzungen der AuftraggeberInnen. Ein hoher Wettbewerb insbesondere bei kleineren Projekten steht einem zunehmenden Zeitdruck in der Verwaltung sowie budgetären Beschränkungen gegenüber. Einige AuftraggeberInnen bestätigen, dass Anbieter aus anderen Ländern selten sind und nur bei größeren Projekten eine gewisse Rolle spielen. Auch die Professionalität der österreichischen Anbieter wird bestätigt. Ihre Kompetenz wird als „sehr hoch“ (50 %) bzw. „hoch“ (50 %) eingeschätzt (jeweils vier von acht). Es findet sich auch die Wahrnehmung, dass bei Anbietern aus anderen Ländern keine höhere Qualität als bei österreichischen Anbietern festzustellen ist⁸.

3.4 VERGABEPRAaxis

Während AuftraggeberInnen im Jahr 2016 im Schnitt 2,2 Teilnahmeanträge bzw. Angebote für einen zu vergebenden Evaluierungsauftrag erhielten, stieg die Zahl 2017 auf vier. 2018 ging der Wert wieder auf 3,5 zurück. Über diesen Zeitraum hinweg waren in den verschiedenen ausschreibenden Stellen jedoch teils große Schwankungen festzustellen. Dabei war in wenigen Fällen lediglich eine Einreichung pro Evaluierungsauftrag zu verzeichnen; in einem anderen Fall fanden acht Einreichungen den Weg zur ausschreibenden Stelle. Dieses Bild wird durch die Sekundärdatenanalyse im Großen und Ganzen bestätigt. Rund 90 % der Teilnahmeanträge bzw. Angebote kamen von österreichischen Einrichtungen.

Im Durchschnitt über alle Befragten hinweg geben diese an, dass Evaluierungen mehrheitlich im Unterschwellenbereich vergeben wurden (vgl. Abb. 3). Dieses Ergebnis wird durch die Sekundärdatenanalyse gestützt. Hier konnten für denselben Zeitraum nur eine Handvoll EU-weit veröffentlichter Ausschreibungen (die also über dem Schwellenwert lagen) identifiziert werden. Da im Unterschwellenbereich angesiedelt, kam es in den allermeisten Fällen (zwei Drittel) zu Direktvergaben, deutlich

weniger wurden in Form von Verhandlungsverfahren mit vorheriger Bekanntmachung (13 %) und offenen Verfahren (11 %) abgewickelt. Der Zuschlagswert, also jener Betrag, für den die Evaluierung schließlich beauftragt wurde, lag laut AuftraggeberInnen in 80 % der beauftragten Fälle in etwa gleichauf mit dem geschätzten Auftragswert.

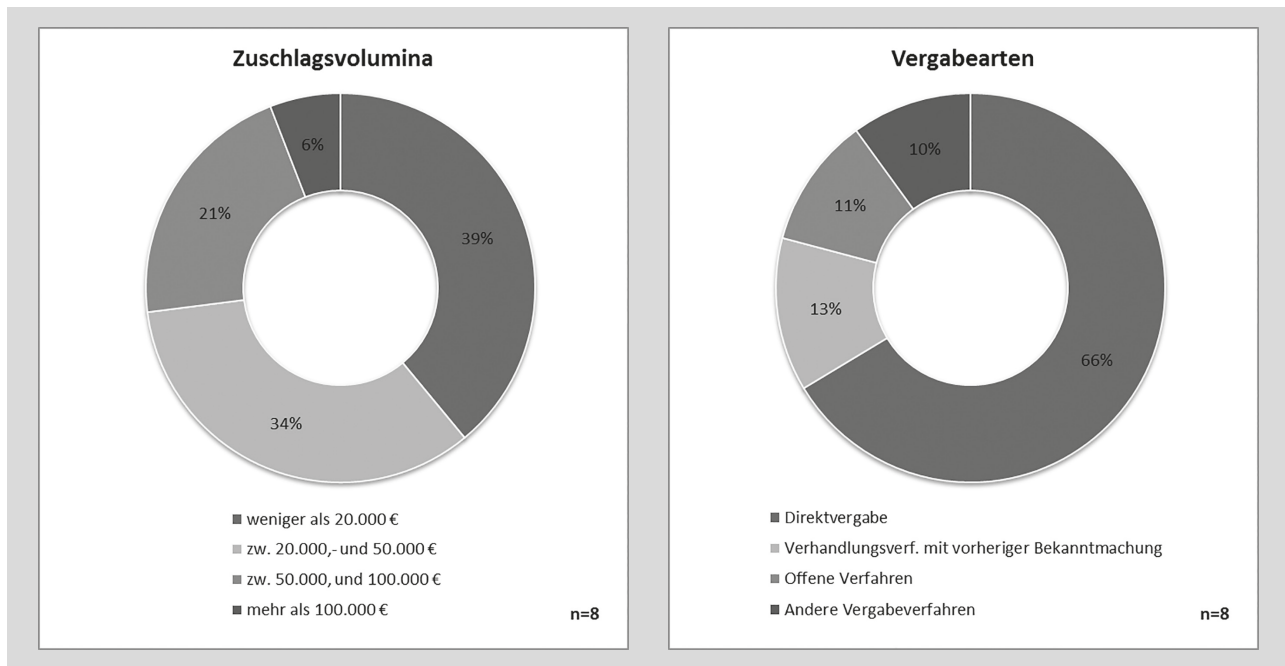


Abbildung 3: Zuschlagsvolumina und Vergabearten, 2016-2018

Quelle: Online-Befragung AuftraggeberInnen 2019. Werte sind arithmetische Mittel.

Gefragt nach den wichtigsten Herausforderungen bei Vergabeprozessen im Inland, nannten die Auftragnehmer am häufigsten sehr breite Anforderungen und detaillierte Fragestellungen sowie die Vorgabe enger budgetärer und zeitlicher Rahmen bei gleichzeitig hohen Anforderungen an methodische Tiefe und Innovation. Auf internationaler Ebene stellen vor allem Rahmenverträge, teils ‚erzwungene‘ Konsortialbildung sowie lange Entscheidungshorizonte Herausforderungen dar. Aus Sicht der AuftraggeberInnen liegen die Herausforderungen im Inland insbesondere darin, das für die Evaluierung verfügbare Budget mit der geplanten Aufgabenstellung in Einklang zu bringen, die Erstellung von Vergabe- und Bewertungskriterien zu definieren und dabei das Vergaberecht einzuhalten.

4. ZUSAMMENFASSUNG UND FAZIT

Die vorliegende Studie liefert erste Einblicke in die Entwicklung des österreichischen Evaluierungsmarktes im FTI-Bereich in Bezug auf die MarktteilnehmerInnen, Wettbewerbssituation und Vergabepaxis. Dass im Zuge der Befragung ein Großteil der relevanten MarktteilnehmerInnen erreicht wurde, spricht für eine hohe Validität der Ergebnisse, welche zudem im Rahmen einer Diskussionsveranstaltung der fteval Anfang Juli 2019 breite Zustimmung fanden.⁹

Die manchmal geäußerten Vermutungen, wonach der österreichische Evaluierungsmarkt von einer geringen Wettbewerbsintensität und einer vergleichsweise kleinen Zahl an AnbieterInnen geprägt sei, konnten im Rahmen dieser Untersuchung nicht bestätigt werden. Vielmehr wird der Markt der FTI-Evaluierungen in Österreich sowohl von AuftraggeberInnen als auch AuftragnehmerInnen als sehr wettbewerbsintensiv beschrieben. Dabei stellt der Markt die Auftragnehmer vor allem durch seine geringe Größe, eine wahrgenommene Konzentration aufseiten der Auftraggeber, den Markteintritt neuer MitbewerberInnen sowie wach-

sende inhaltliche Anforderungen bei geringen Budgets vor große Herausforderungen.

Die Wettbewerbsintensität durch neue Anbieter ist hoch. Obwohl die Zahl der Mitglieder der Plattform fteval in den letzten Jahren gewachsen ist, entfällt ca. ein Viertel aller Aufträge auf Institutionen, die nicht Mitglieder der Plattform sind. In Richtung hoher Wettbewerbsintensität weisen auch die Antworten auf Fragen über die preisliche Wettbewerbsfähigkeit. Ein nicht unbeträchtlicher Teil der MarktteilnehmerInnen bezeichnet die Projekte als oftmals nicht kostendeckend.

Die Untersuchung bestätigt teilweise die Vermutung über den geringen Internationalisierungsgrad des österreichischen Evaluierungsmarktes. Es ist zu vermuten, dass der österreichische Evaluierungsmarkt mit seiner insgesamt relativ geringen Größe, den geringen Projektgrößen, und der wachsenden Zahl an Mitbewerbern für ausländische Anbieter nur begrenzt attraktiv ist. Hinzu kommt eventuell das Spezifikum von Deutsch als Berichtssprache. Die Nachfrage ist von Direktvergaben gekennzeichnet; diese sind oft weniger transparent als andere Vergabeformen. Anbieter aus anderen Ländern sind kaum im Lead von Konsortien zu finden, und auch unter den Konsortialmitgliedern schwach vertreten.

Auf der Seite der von den AuftraggeberInnen wahrgenommenen Evaluierungskompetenz der österreichischen Anbieter dürfte es kein Problem geben: Diese wird durchwegs – auch im Vergleich mit ausländischen Anbietern – als hoch bzw. sehr hoch eingeschätzt. Zudem sind österreichische Anbieter von Evaluierungen vielfach auch auf internationaler Ebene aktiv.

Insgesamt weisen die vorliegenden Befunde auf einen wettbewerbsintensiven FTI-Evaluierungsmarkt in Österreich hin, der aber zum allergrößten Teil von einer Vielzahl österreichischer AnbieterInnen bedient wird, denen eine hohe Qualität bescheinigt wird. Gleichzeitig weist er einige Charakteristika auf, die sich negativ auf eine qualitative Entwicklung auswirken können: geringe Projektgrößen, hoher Preisdruck und ein hoher Anteil an Direktvergaben könnten solche potentiell retardierende Faktoren sein. Versuche, den Markt für FTI-Evaluierungen qualitativ weiter zu entwickeln, sollten hier ansetzen.

Abschließend ist nochmals zu betonen, dass die gegenständliche Studie nicht alle Dimensionen des FTI-Evaluierungsmarktes, sondern nur den klar abgegrenzten Bereich der Vergabe öffentlicher Aufträge für wissenschaftliche Dienst- und Forschungsleistungen abdecken konnte. Darüber hinausgehende Fragen, etwa zur Bedeutung von Studien, Analysen und wissenschaftlichen Begleitaktivitäten, die nicht Evaluierungen entsprechend der hier zugrunde gelegten Definition umfassen, bedürfen näherer Untersuchungen. Gleichwohl hoffen wir, mit dieser Studie eine (erste) empirische Basis für eine fundierte Diskussion gelegt zu haben. Eine anzustrebende wichtige Erweiterung der Analyse wäre ein Vergleich des österreichischen Marktes mit anderen Ländern.

5. REFERENZEN

- Aiginger, K., Falk, R. und Reinstaller A.** (2009): Evaluation of government funding in RTDI from a systems perspective in Austria. Synthesis Report, Vienna.
- Biegelbauer, P.** (2013). *Wie lernt die Politik? Lernen aus Erfahrung in Politik und Verwaltung.* Springer-Verlag.
- BMWF, BMVIT und BMDW** (2018): Österreichischer Forschungs- und Technologiebericht 2018. Bericht der Bundesregierung an den Nationalrat gem. § 8 (2) FOG über die Lage und Bedürfnisse von Forschung, Technologie und Innovation in Österreich, Wien.
- BMFW und BMVIT** (2017): Österreichischer Forschungs- und Technologiebericht 2017. Lagebericht gem. § 8 (1) über die aus Bundesmitteln geförderte Forschung, Technologie und Innovation in Österreich, Wien.
- Böttcher, W., Kerlen, C., Maats, P., Schwab, O., & Sheikh, S.** (2014). *Evaluation in Deutschland und Österreich.* Münster: Waxmann.
- DeGEval Standards** (2016): DeGEval – Gesellschaft für Evaluation. Standards für Evaluation. Erste Revision auf Basis der Fassung 2002. https://www.degeval.org/fileadmin/Publikationen/DeGEval_Standards_fuer_Evaluation_-_Erste_Revision__2016_.pdf
- Dinges, M. und Schmidmayer, J.** (2010): Country report: Austria., in: Edler, J., Cunningham, P., Gök, A., Rigby, J., Amanatidou, E., Garefi, I., Bühner, S., Dinges, M., Berger, M., Schmid-mayer, J. und Guy, K. (2010): *Understanding evaluation of innovation policy in Europe*, in: INNO-Appraisal final report, Manchester, 243–264
- fteval Standards** (2019): Standards der Evaluierung in der Forschungs-, Technologie- und Innovationspolitik. https://repository.fteval.at/386/7/standards_DT_WEB_08042019.pdf
- Handler, H.** (2005): Öffentliches Auftragswesen und Wettbewerb im gesamtwirtschaftlichen Zusammenhang; in: Michael Sachs (Hrsg.), *Schwerpunkte zum BVergG 2006: Ökonomische und rechtliche Aspekte*, Wien (Manz) 2005: 27-58.
- Landsteiner, G.** (2015): Nützlichkeit und Nutzen der Programmevaluationen im Bereich der österreichischen FTI-Politik. *Metaevaluation der Programmevaluationen 2003–2014*
- Reiner, D. und Smoliner S.** (2012): Outputorientierte Evaluierung öffentlich geförderter FTI-Programme. Studie im Auftrag des Bundesministeriums für Verkehr, Innovation und Technologie, Wien
- Streicher, J.** (2017): *Evaluations, Actors and Institutions. The Case of Research, Technology and Innovation Policy in Austria*
- Tsipouri, L und Sidiropoulos, N.** (2014): RTDI evaluation culture in the EVAL-INNO countries. In: Zupan I, Gajdusek MF and Marinkovic I (Hsg.) *Fostering Evaluation Competencies in Research, Technology and Innovation – the EVAL-INNO experience*, S. 11-20. LIT Verlag.
- Warta, K. und Philipp, S.** (2016): Was bringt die Plattform fteval und zu was hat sie es ge-bracht? Rückblicke und Ergebnisse einer Mitgliederbefragung, *The fteval Journal for Research and Technology Policy Evaluation*, Ausgabe 41, 5–14

6. ABBILDUNGSVERZEICHNIS

Abbildung 1:	Anzahl an identifizierten Evaluierungen im Zeitraum 2009-2018	76
Abbildung 2:	Einschätzung der Wettbewerbssituation, 2016-2018	78
Abbildung 3:	Zuschlagsvolumina und Vergabearten, 2016-2018	79

7. TABELLENVERZEICHNIS

Tabelle 1:	Verwendete Quellen von Sekundärdaten	73
Tabelle 2:	Evaluierungen auf nationaler Ebene, 2002-2007 sowie 2016-2018	76

AUTOREN

JÜRGEN STREICHER

Joanneum Research
Sensengasse 1, 1090 Wien
E: juergen.streicher@joanneum.at

WOLFGANG POLT

Joanneum Research
Sensengasse 1, 1090 Wien
E: wolfgang.polt@joanneum.at

MAXIMILIAN UNGER

Joanneum Research
Sensengasse 1, 1090 Vienna
E: maximilian.unger@joanneum.at

KEYWORDS:

Evaluierungsmarkt; FTI-Evaluierung; Österreich

NOW AVAILABLE!

LATEST EDITION OF THE

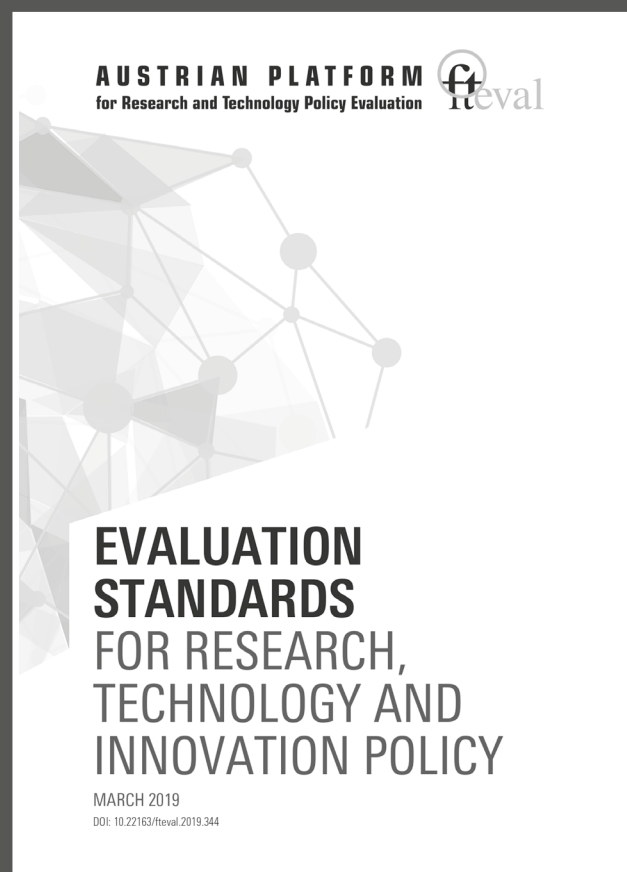
'EVALUATION STANDARDS FOR RESEARCH, TECHNOLOGY AND INNOVATION POLICY'

OF THE AUSTRIAN PLATFORM FOR RESEARCH
AND TECHNOLOGY POLICY EVALUATION



Download in German:

fteval.at/content/home/standards/fteval_standards/



Download in English:

fteval.at/content/home/standards/fteval_standards/index.jsp?langId=2

Do you want print copies?

Please send an eMail to office@fteval.at

FTEVAL JOURNAL FOR RESEARCH AND TECHNOLOGY POLICY EVALUATION



... RTI POLICY EVALUATION AT THE INTERFACE BETWEEN PRACTICE AND ACADEMIA ...

The 'fteval Journal for Research and Technology Policy Evaluation' contributes in a quality assured way to exchange between various stakeholder groups in the area of science, research, technology and innovation policy evaluation. It addresses policy-makers, practitioners, evaluators and the academic community. Thematic editions alternate with thematically open issues. By now 50 issues have been published, which can be accessed and downloaded from the fteval-repository and the fteval homepage. The fteval Journal is open access. The journal and each published paper receive a separate DOI. All paper contributions are quality controlled, but not subject to peer-review in the strict academic sense.

Researchers and practitioners from the RTI policy domain are invited to send contributions to the editor of the 'fteval Journal for Research and Technology Policy Evaluation' (office@fteval.at). The contributions can be submitted either in German or English.

If you want to receive the 'fteval Journal for Research and Technology Policy Evaluation' free of charge, please fill out this form and send it to office@fteval.at or drop it into the infobox at one of our conferences and events.

first name

last name

institution

address

postal code

city/place

country

e-mail

By filling in this information you agree to* receive the fteval Journal by mail (free of charge)
 * to receive the electronic newsletter of fteval, which informs about pertinent issues of RTI evaluation in Austria and the EU via email (such as calls for contributions for the next fteval Journal) and thus you entitle the fteval to process your indicated personal data for the purpose of transmission of information to you.
 You can revoke consent at any time by informing us via eMail (office@fteval.at).





fteval JOURNAL for Research and Technology Policy Evaluation

ISSUE 51

Europe

CALL FOR PAPERS JOURNAL #51

The upcoming fteval journal's topic is **Europe**. We welcome both, academic and practice-oriented contributions, dealing with the evaluation of RTI policies with a European dimension. Articles may range from specific results of evaluation studies, to methodological, organisational or theoretical contributions. Please share your ideas for contributions as **abstracts until April 17, 2020**. The deadline for the **submission of final papers is July 15, 2020**.

If you have any questions about the journal, we remain at your disposal at office@fteval.at.

EDITORIAL BOARD

Rebecca Allinson, **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 (chief editor), **Austrian Platform for Research and Technology Policy Evaluation**; 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, 1150 Vienna, Austria

T +43 1 495 04 42 - 33

F +43 1 495 04 42 - 40

E office@fteval.at

W www.fteval.at

DESIGN

W carotte.at

PRINT

Print Alliance

Gedruckt auf PEFC™-zertifiziertem Papier.

AUSTRIAN PLATFORM FOR RESEARCH AND TECHNOLOGY POLICY EVALUATION

members:

Austrian Federal Ministry of Education, Science and Research (BMBWF), Austrian Federal Ministry for Digital and Economic Affairs (BMDW), Austrian Federal Ministry for Climate Action, Environment, Energy, Mobility, Innovation and Technology (BMK), Austrian Cooperative Research (ACR), Austrian Council for Research and Technology Development, Austrian Institute of Technology (AIT), AQ Austria – Agency for Quality Assurance and Accreditation Austria, Austria Wirtschaftsservice (AWS), Christian Doppler Research Association (CDG), convelop Cooperative Knowledge Design GmbH, Austrian Research Promotion Agency (FFG), Austrian Science Fund (FWF), Institute for Advanced Studies (IHS), Industriewissenschaftliches Institut (IWI), Joanneum Research Forschungsgesellschaft mbH, Austrian Institute for SME Research (KMU Forschung Austria), Ludwig Boltzmann Society (LBG), Austrian Academy of Sciences (OEAW), Technopolis Austria, Vienna Business Agency – A service offered by the City of Vienna, Austrian Institute of Economic Research (WIFO), WPZ Research GmbH, Vienna Science and Technology Fund (WWTF), Centre for Social Innovation (ZSI)

ISSUE №50

AUSTRIAN PLATFORM
for Research and Technology Policy Evaluation

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

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 - 33
F +43 1 495 04 42 - 40
E office@fteval.at
W www.fteval.at

ZVR-Zahl: 937261837
ISSN-Nr. 1726-6629
DOI: 10.22163/fteval.2020.462
© Vienna 2020