fteval JOURNAL
for Research and
Technology Policy Evaluation

July 2025, Vol. 57, pp. e3, 1-25
DOI: 10.22163/fteval.2025.695

THE ROLE OF CHANGE AGENT CHARACTERISTICS IN RESEARCH ASSESSMENT:

EXPERIENCES FROM AN EVALUATION OF A RESEARCH FELLOWSHIP PROGRAM IN GERMANY

HENDRIK BERGHÄUSER
DOI: 10.22163/FTEVAL.2025.695

ABSTRACT

Research funders are increasingly faced with the challenge to design funding programs generating both scientific and social impact, which in turn places new demands on research assessment processes. Using an evaluation of a german research funding program targeting scientists from the Global South, this paper examines how different funding goals and selection criteria - in this case research excellence on the one hand and social or development relevance on the other - can be combined in research assessment and to what extent change agent qualities of the applying researchers can be integrated into the assessment process in order to increase the chances for social impact. Drawing on interviews, a survey of funded researchers as well as a survey of a control group of non-funded scientists the study shows that funded researchers indeed, have much stronger change agent attitudes than scientists in the non-funded control group. Differences with regard to concrete change agent actions, on the other hand, cannot be determined. All in all, change agent characteristics represent a promising element for research assessment procedures, but the selection criterion suffers from an insufficient definition and conception, especially for the science sector.

Keywords: Research assessment, Social impact, Change agents, Global South

1. INTRODUCTION

As research increasingly seeks to contribute to solutions for social problems, with strong emphasis on transfer activities and social engagement (Belcher et al. 2016; Bornmann 2013), research funders are faced with the challenge to design funding programs which generate both social and scientific impact. This implicates a need for research assessment processes that take into account and deal with tradeoffs in both research excellence and the potential of social impact, especially when diverse actors and researchers outside mainstream circles are addressed (Kraemer-Mbula et al. 2020; Ferretti et al. 2018). This paper provides insights on an evaluation of a research fellowship program that aims to enable excellent scientists from developing and emerging countries to spend a research period in Germany. The fellowship program wants to recognize the relevance of researchers from these countries in achieving the 17 UN Sustainable Development Goals (SDGs) (UN General Assembly 2015). Thus, classic indicators of research excellence such as key publications, scientific impact (h-index) etc. fall short or need to be contextualized, given the heterogeneity of researchers from 137 eligible countries and the partly difficult conditions in the science systems in the Global South (Tijssen and Winnink 2022). At the same time, the notion of "social impact" or "contributions to SDGs" appear abstract and elude standardized, indicator-based objective evaluation. Therefore, in addition to scientific merits, the research assessment jury of this funding program also looks for change agent characteristics among the applicants (Caldwell 2005; Hutton 1994; Ottaway 1983). This is linked to the assumption that certain personal characteristics increase the likelihood that the results of the funded research projects will eventually find their way into application in the home countries of the researchers and thus generate a developmentally relevant impact.

The focus of this paper is therefore twofold: First, it examines and discusses possibilities and limits when linking two different funding objectives - research excellence and social or developmental impact - as well as potential conflicting goals. Second, the paper looks at the potential of change agent characteristics of the applicants as a criterion for research assessment in individual funding programs.

The evaluation object is the Georg Forster Research Fellowship of the Alexander von Humboldt Foundation, which is funded by the Federal Ministry for Economic Cooperation and Development in Germany. The selected researchers receive a monthly sponsorship for a research stay in Germany of

up to two years and are hosted by a german scientific institution during their research period.[1] Decisions to award fellowships are based on the assessment of the applicant's academic qualification, their contribution to development-related research as well as their future potential.

2. THE ROLE OF CHANGE AGENT CHARACTERISTICS IN RESEARCH ASSESSMENT

While studies on the contribution of success in academic research can be traced back to over a half a century (Merton 1957; Zuckerman and Merton 1971; Hagstrom 1975) debates about research assessment processes intensified in the 1980s, when the growth in public research funding led to a growing workforce and a stronger need for clear and transparent distribution of scarce financial resources for funding allocation, hiring, tenure and promotion became obvious (Alberts et al. 2014; Bonn and Bouter 2023). As a consequence, publication metrics, in particular publication counts, citations, the H-index as well as the journal impact factors started being used in research assessment in order to provide a greater sense of objectivity than traditional forms of peer-review qualitative assessment (Gingras 2014; Bonn and Bouter 2023). However, due to the strong focus on these mostly quantitative metrics, research assessment has come under increasing pressure. In particular, the underlying concept of research excellence and the application of too narrow criteria and indicators of research quality was criticized. According to this critique, the strong focus on these metrics or the narrowing of quantitative criteria and indicators reduces the diversity of research missions and purposes, leading researchers to adopt similar strategic priorities or to focus on lower-risk, incremental work. Furthermore, the systemic biases against those who do not meet these narrowed indicators of quality or impact, or who do not conform to certain career pathways, reduce the diversity and representative legitimacy of the research community (Curry et al. 2020; Moed 2020; Tijssen and Winnink 2022). In addition, the application of these metrics, combined with an increasing competition for research funding might distort incentives, create unsustainable pressures on researchers and thus lead to unethical behavior (Edwards and Roy 2017; Moher et al. 2018). Therefore, efforts were coalescing around the idea of a more responsible approach of

1 For more information about the funding program, see: https://www.humboldt-foundation.de/en/apply/sponsorship-programmes/georg-forster-research-fellowship.

metrics and its use that lead to new initiatives in research assessment like the Declaration on Research Assessment (DORA) in 2012 (DORA 2021), the Leiden Manifesto for research metrics (Hicks et al. 2015) or the Metrics Tide (Wilsdon et al. 2015) that recognized the need to improve the ways in which researchers and research outputs are evaluated.

At the same time, an increased focus on directionality and mission-orientation can be observed in science policy. This is accompanied by expectations of research to develop concrete solutions for societal challenges (Boon and Edler 2018; Mazzucato 2018), leading scholars to conceptualize impacts beyond academia and the contribution of science to addressing current or future social, environmental, economic, policy and other needs or problems (D'Este et al. 2018; Reale et al. 2017) as well as to develop frameworks to evaluate such impacts (Kok and Schuit 2012; Joly et al. 2015; Smit and Hessels 2021; Spaapen and van Drooge 2011; Matt et al. 2017; Donovan and Hanney 2011; Lauronen 2020).

While research assessment procedures are under pressure, the demands placed on researchers applying for research funding have also changed, especially in funding programmes targeting both social and academic impact. Since funded researchers are expected to promote not only scientific but also social impacts through their research work, personal characteristics, which can generally be described as change agent characteristics, are increasingly coming to the fore in research assessment alongside aspects of research excellence. The first comprehensive definition of 'change agents' was developed by Beckhard (1969) who describes change agents as people either inside or outside of an organization who provide technical, specialist or consulting assistance in the management of a change effort (Beckhard 1969). Roger and Shoemaker define change agents as professionals who influence innovation decision in a desirable direction (Rogers and Shoemaker 1971). They bring about purposeful transformation and help people to change the way they think - changing the norms and changing the organization's systems and processes (Hutton 1994). Moran and Brightman (2000) look at the necessary characteristics of change agents and find out that persons need to have necessary skills to initiate and manage change processes, e.g. in leadership, creativity and problem-solving (Moran and Brightman 2000). More recent studies underline the personal skills needed for change agency like openness, empathy, energy and networking Lunenburg (Lunenburg 2010). As a consequence, researchers can also act as change agents and contribute to social impact through promoting public values and knowledge transfer activities or by focusing on the social relevance of research (Bornmann 2013).

The concept of change agents could therefore also link the concept of social impact and research excellence. Tijssen and Winnink (2022) differentiate in this context between 'global excellence' and 'local excellence'. Whereas global excellence is primarily determined by acknowledged scientific visibility and (partially) measurable reputation within the international research community, local excellence instead relates to the utilisation of knowledge and know-how among non-scientific users and local communities. Nevertheless, the concept of change agents in science remains vague and ill-defined.

3. METHODOLOGICAL APPROACH

The methodological approach used here is based on a triangulation of different quantitative and qualitative as well as reactive and non-reactive empirical methods used for the evaluation of the mentioned funding programs. The analysis relies on participant observation of two meetings of the research assessment jury based on an observation guide, five interviews with program managers and representatives of the research assessment jury, an online survey of all research fellows who have received program funding between 2010 to 2020 and a control group survey with non-funded scientists from developing and emerging countries with comparable researcher characteristics identified by using bibliometric methods.

4. PROGRAM STAKEHOLDER INTERVIEWS AND PARTICIPANT OBSERVATION OF RESEARCH ASSESSMENT JURY MEETINGS

The evaluation process started with a group interview with several program managers from the funding agency in Spring 2022. The goal of this initial interview was to obtain more detailed information about the funding program, to discuss the overall goals of the program and to better understand the underlying program logic. In addition, the procedures of the triannual research assessment jury meeting and the criteria for selecting applicants for the research fellowship were discussed during this interview. Consequently, this group interview provided the basis for the subsequent planned participant observation of two research assessment jury meetings.

After the interview and an analysis of funding guidelines and further program information for applicants, the preparation for the participant observation of the

research assessment jury meetings started. The participant observation used in this case referred to an unstructured procedure with an explorative objective of recording background conditions (Kawulich 2005). Specifically, two research assessment jury meetings were observed (in June and October 2022). As a prerequisite, all stakeholders involved (program managers and jury members) were informed about the process of the participant observation. Consent was obtained from all participants. The participant observation was then conducted in an open and non-active participatory manner, i.e. the members of the research assessment jury were aware of the participation of the evaluators, yet the evaluators did not actively participate in the discussion but merely observed the process instead. The observation followed a previously developed observation scheme in which the objects of observation and the tasks are concretized. The strength of this method is that it allows the observation of research assessment and selection processes and thus actual behavior regarding the implementation of the two selection criteria. The observation guide was derived from the previous interview and the written selection criteria:

Observations regarding assessment criteria "research excellence"	Observations regarding assessment criteria "social / developmental relevance"
▪ Observations regarding the topic of scientific career and scientific achievements (mobility, determination, breadth of expertise, scientific productivity)	▪ Observations regarding the relevance of the research projects for the further development of the country or region of origin (including the 17 SDGs of the UN)
▪ Observations regarding the quality of the key publications name in the application (originality, degree of innovation, own contribution in the case of multi-author publications)	▪ Observations regarding change agent characteristics of the applicant or possible multiplier effect in research, teaching, science management and development-relevant processes and activities outside academia
▪ Observations regarding the originality and innovative potential of the proposed research project (significance for the further development of the field, convincing choice of scientific methods, possibilities for further scientific development, feasibility, etc.)	▪ Observations regarding further development-relevant aspects (for example gender-specific aspects, special need for support with regard to other specific region or the concrete research area)
▪ In the case of postdocs, how is the scientific potential of the applicant assessed (further scientific development, career prospects etc.)? In the case of experienced researchers: How is the stand-alone scientific profile of the researcher evaluated?	

In addition to operationalizing the two selection criteria – research excellence and social / developmental relevance – the participating observers were also looking for "soft" aspects regarding the discussion dynamics in particular:

- Dominant actors (jury members dominating the discussion beyond their role),
- Gender-perspective (are all participants fairly engaged in the discussions, regardless of their socio-demographic characteristics? Is there any evidence for homosocial biases, i.e. men favoring men or women favoring women?),
- Balance of power between different actors within the research assessment jury meetings (e.g. how passive are the funders vis-à-vis the experts?),
- How are controversial decisions handled? How are conflicts resolved?

As the research assessment jury meetings took place in a hybrid format, both meetings were observed by two participants, one taking part virtually and one taking part physically. In this way, the authors were able to evaluate how the two main selection criteria – research excellence and social / developmental relevance of the applying researchers and their proposed research projects were being assessed and how these two selection criteria were connected in the discussion within the jury.

The two observers recorded their observations of the two half-day research assessment jury meetings in a result protocol. The results were then compared and discussed. Divergent perceptions and observations were not included in further analysis. Subsequently, the results logs were coded and analyzed using MAXQDA, a qualitative content analysis software.

After the participant observation, four interviews were conducted with members of the research assessment jury. The interviews were semi-standardized and based on an interview guide. The goal of the interviews was to find out about the individual understanding and interpretation of the selection criteria, as well as to mirror observations of the jury assessment meetings (Laudel and Gläser, 2007). Accordingly, the interviews addressed the perception and linkage of research excellence and social or developmental relevance, including change agent characteristics, in the selection process. When selecting the jury members for interviews, we paid attention for a high diversity with regard to disciplinary orientation (e.g. humanities, natural sciences, engineering), gender and region-specific knowledge, which were of great importance in the evaluation of the applicants from the global south. The interview partners were selected after the jury meetings. Therefore, also

the behaviour of the reviewers during the jury meetings was also taken into account. For example, among those selected were reviewers who highlighted or critically questioned certain selection criteria (e.g. 'excellence' or 'social / developmental relevance') during the discussion of the applications. The interviews were recorded and transcribed. Subsequently, the transcripts were also coded and analyzed using MAXQDA software.

4.1 ONLINE-SURVEY AMONG FUNDED RESEARCH FELLOWS

In fall 2022, preparations began for the quantitative analyses, in particular the bibliometric analyses as well as the online surveys of the selected and funded research fellows and of the control group. On December 5, 2022, the survey of the funded fellows was launched, using the EFS survey tool by Questback. A total of 754 persons were invited to take part, of which 20 persons could not be reached for technical reasons (final adjusted total gross sample: 734). On December 14, a reminder mail was sent out, in which all those persons who had not yet participated by that time or who had dropped out of the survey were contacted again. A second and final reminder campaign was launched on January 9, 2023. The survey was closed on January 24. The results dataset was then downloaded and cleaned. In the end, 505 valid responses were counted, which corresponds to a substantial participation rate of 68.8%.

4.2 CONTROL GROUP SURVEY AMONG NON-FUNDED BIBLIOMETRIC TWINS

In the evaluation, bibliometric analyses were mainly used to investigate the publication performance of the funded fellows (before and after the funded research stay) and to measure the development of the fellows' international scientific collaboration (through co-publications). In addition, bibliometrics was also used as an approach to generate a control group to be surveyed subsequently. For the bibliometric formation of the control group, Scopus database was used. For each funded research fellow, one or more "twins" were bibliometrically generated by searching for a scientist with similar characteristics. There characteristics were:

- Scientific discipline: Here the journals, in which the funded fellow mainly publishes served as a reference,
- Gender (male / female): this is generated from the first name of the scientist,
- Scientific age: This is calculated based on the first scientific publication covered in Scopus,

- The country / region of origin: a minimum number of publications in the years prior to the funding year (of the funded fellow) in the funded fellow's home countries or regions,

Publication performance: A similar publication output of a researcher and citation rate at the time the fellow is funded. In a first step, however, the funded fellows needed to be identified in Scopus. Of the 809 funded research fellows, 651 could be clearly identified by name and e-mail addresses (80.5%).

For these 651 researchers, "twins" were then searched for in Scopus using the above defined criteria. In order to obtain a larger sample for the control group survey, multiples were also counted (i.e. if there was more than one bibliometric twin for a fellow). However, this also means that some research fellows had a stronger weight in generating the comparison group than others. As a result, complete representativeness was not achieved. But this was taken into account when comparing the survey data. Nevertheless, these methodological limitations were accepted in order to achieve a sufficiently large sample that promised a solid response in an online survey. In the end, 1918 researchers were identified as potential twins for the control group survey. The following table compares the two groups surveyed in terms of key characteristics.

Research fellows		Control group	
Gender		**Gender**	
Male	71.2%	Male	67.5%
Female	27.5%	Female	29.8%
Diverse / other	0.4%	Diverse / other	0.4%
n/a	0.8%	n/a	2.4%
Discipline		**Discipline**	
Agriculture, forestry, horticulture and veterinary medicine	9.6%	Agriculture, forestry, horticulture and veterinary medicine	11.4%
Biology	16.7%	Biology	13.3%
Chemistry	16.2%	Chemistry	7.5%
Humanities	15.6%	Humanities	16.1%
Geosciences	3.8%	Geosciences	2.7%
Engineering sciences	11.3%	Engineering sciences	10.6%
Mathematics	2.8%	Mathematics	1.6%
Medicine	3.8%	Medicine	5.5%
Physical sciences	3.8%	Physical sciences	3.5%
Social and behavioural Sciences	10.3%	Social and behavioural Sciences	16.5%
Other	6.0%	Other	11.4%
Scientific age (here: date of doctorate degree)		**Scientific age (here: date of doctorate degree)**	
Before 2000	1.2%	Before 2000	8.7%
2000 – 2004	9.3%	2000 – 2004	6.2%
2005 – 2009	23.0%	2005 – 2009	16.6%
2010 – 2014	36.2%	2010 – 2014	29.1%
2015 – 2019	20.6%	2015 – 2019	15.2%
2020 and later	0.6%	2020 and later	8.7%
n/a	9.1%	n/a	15.6%
Country of origin		**Country of origin**	
651 research fellows from 62 countries in the Global South were identified in Scopus		Comparable twins were identified from 59 of 62 countries.	
Publication performance		**Publication performance**	
individual		Identical number of pub. + / - 3 publications + / - 5 publications	58.2% 26.6% 15.2%

Some of the key characteristics, for identifying bibliometric twins were validated in both surveys and could thus be determined even more precisely. The participants of the surveys were asked for examples about their scientific discipline, their gender and the year of their doctorate's degree. In spite of the mentioned methodological limitations, the two groups are very similar in terms of key characteristics, as table 2 shows. There are slight differences in the distribution of scientific age and publication performance.

The control group survey started on March 1, 2023. Of the 1918 identified researchers, 1753 could be reached. An initial reminder campaign was launched on March 12 and a second and final one on March 26. On April 3, 2023, the control group survey was closed, and the results dataset was cleaned according to the same procedure as the dataset of the fellowship survey. The final number of participants was 189, which represents a participation rate of 16.5%.

In order to avoid a selection bias, the control group survey was intentionally framed rather broadly in terms of content. In the invitation email, the aim of the study was described as basically asking about the motives and attitudes of researchers in the Global South. In this context, corresponding questions on the role of science in society and about change agent characteristics were asked and analysed. Therefore, a possible selection bias can be largely excluded. The large difference in the participation rate between both surveys is, primarily due to the close relationship between the funded researchers and the funding agency and the explicit reference in the survey of funded research fellows, that the survey is related to the evaluation of the funding programme.

5. RESULTS

5.1 PROGRAM STAKEHOLDER INTERVIEWS AND PARTICIPANT OBSERVATION OF RESEARCH ASSESSMENT JURY MEETINGS

The jury meeting was preceded by an external, written assessment of all applications by at least two reviewers. In the jury meetings, only those applications were presented and discussed for which the external reviews did not provide a clear picture. Applications with two poor external reviews for example were already excluded before the jury session. The same applies to excellent applications that received two very good external reviews. These clear funding cases ("F-cases") were also not discussed at the jury meeting. These candidates received funding without an additional discussion

and jury vote. Consequently, the assessment jury only dealt with the so-called discussion cases ("D-cases") which, however, make up the majority of applications. The research assessment jury meetings proceeded in a highly structured manner. Each jury member presented the applications that were assigned to them based on their professional profile and regional expertise. Within a few minutes, the respective applicant and the proposed research project were briefly presented and discussed. In doing so, the jury member gave his or her assessment about the research excellence and the developmental relevance and made a final recommendation for or against funding. In the assessments, reference was also made to the external review reports that were available to all jury members. Afterwards, there was a short question and answer session before the next application was presented. Voting was done in secret at the end of the meeting (each jury member could distribute 0,1 or 2 points for each candidate).

During the presentation and discussion of the applications, a rather benevolent review could be observed overall. The jury members predominantly expressed respect and appreciation for the candidates ("very good candidate", "very interesting project", "very determined candidate", "has high potential", "very impressive" etc.).

In addition, the candidates were presented and assessed in comprehensive form. The positive aspects and negative impressions were weighed up. In the discussion, one single criterion was rarely decisive for the assessment. Moreover, all selection criteria (e.g. academic career, scientific productivity, key publications or developmental relevance) were considered with regard to the individual context. Individual characteristics such as age, gender, origin as well as private life circumstances were explicitly taken into account in the evaluation ("The candidate is already somewhat older. But she has raised three children in Africa and earned a doctorate in Japan. That is impressive"). The individual context conditions in the home countries of the applicants were also considered in the assessment ("the publication performance of this candidate is not outstanding, but it must be taken into account that the candidate, as is common in this country, was employed as a lecturer for a long time, which meant a high teaching load and little room for research and publication").

However, it remains unclear, how exactly the different levels of development in the applicants' home countries were weighted. At the opening of the meeting, the funding authority pointed out the applications from scientists from least developed countries (LDCs) and asked for a particularly favorable evaluation. In the subsequent interviews with jury members after the jury

meetings, it was said that applicants are not directly compared with each other. Nevertheless, a very good candidate from a very poor country may receive a higher rating than an equivalent candidate from a more developed country, since the performances are to be weighted differently with regard to the different conditions in the home countries. However, in some cases, especially in the case of candidates from the least developed countries, it was discussed whether and to what extent the candidates would be able to continue their work in his or her home country after a research stay in Germany. In one case, this was doubted. It was argued that there were no appropriate infrastructures in the home country to continue this research or even to establish it in the first place. In this case, funding was not recommended, since the probability of generating social impact through the fellowship was considered rather low and the risk of brain drain through the funded research stay was considered to be high. The discussion made it clear that academic expertise in each discipline coupled with regional knowledge are necessary in order to be able to thoroughly and fairly evaluate a candidate and to contextualize individual achievements ("All the applicants' key publications are in Spanish. But we have to consider that in [this country] science is mainly published in Spanish and addresses local and regional scientific audiences first"). In assessing the potential of the applicants, particular consideration was given to their academic performance over the course of their academic career ("his curve goes steeply upwards" or "In the beginning it was a little less, but now she is very active and publishes a lot"). In addition, a higher age (e.g. 60 years or older) tended to be evaluated rather negatively, since the chances regarding scientific potential or social impact were estimated to be rather low.

Developmental and/or social relevance counts, at least formally, more than scientific excellence in this funding program. In fact, developmental and/or social relevance was evaluated first in the selection process before scientific excellence was evaluated. If no developmental relevance was seen, or if it was deemed to be too low, the application was immediately sorted out and not admitted to the final vote at the end of the session. This procedure also describes how the two selection criteria - scientific excellence on the one hand and developmental and/or social relevance on the other hand - are linked to each other in this funding program: Developmental and/or social relevance is basically the prerequisite for receiving any chance of funding at all. The two selection criteria are not weighed against each other.

At the same time, the assessment of developmental and/or social relevance often remained very vague. Often, the justification for the assessment was not further elaborated ("the proposed project is clearly developmentally

relevant" or "the developmental relevance is given"). Sometimes it remained unclear what the developmental relevance refers to, e.g. which regional focus is considered (developmentally relevant for a specific region, country, continent or even reference to global challenges), or which of the SDGs are addressed or for which social areas or social groups possible impacts are foreseeable. Discussions about developmental relevance mostly arose among more basic research-oriented projects, as well as among applications from social scientists and humanities scholars. Here, members of the research assessment jury more often asked about the development relevance of the research project than in the case of applied-oriented research projects that are more associated with general development cooperation, like projects dealing with micro plastics or the contamination of drinking water. In this context, there were also occasional lively discussions and appeals from jury members not to focus solely on application-oriented research projects, even if the developmental relevance may be more obvious here at first glance. The relevance of basic research, which might not have an immediate impact on society but can contribute to social impacts eventually, was then frequently emphasized. Discussed examples included projects that deal with specific mathematical or physical models that can provide contributions for subsequent weather forecasts or projects in biology for the classification of plant or animal species, in order to obtain contributions for a better understanding of biodiversity in a particular region. All in all, the impression was that developmental relevance needed to be justified more strongly in the case of basic research-oriented projects or in the case of social sciences and humanities. This tension was confirmed in the subsequent interviews with jury members: "We always get to that point with theorists regarding their developmental relevance" (Interview no. 3).

Overall, a very broad understanding of developmental relevance could be observed. According to the jury, the vast majority of the proposed research projects of the candidates were assessed as developmentally relevant to some degree. In fact, only a few applicants were denied on the grounds of the developmental and/or social relevance of their proposed research projects. Consequently, development relevance seems more like a soft selection criterion.

A special aspect in the evaluation of developmental relevance are the change agent characteristics. These too, have not been defined in any further detail at any point. This usually involved a special commitment to science and/or society. Examples given included special commitment to the training of doctoral students and undergraduates, the establishment of study programmes, research areas or entire research institutes or involvement in

various professional societies or committees. In addition, special emphasis was placed on engagement with social actors that are closely related to the applicant's scientific work. This often refers to various transfer activities in which the applicants were engaged. In this context, various examples of knowledge and technology transfer (i.e. candidates planning science-industry cooperation, citizen science projects or candidates consulting policy makers) as well as civic engagement were described (private involvement with an NGO, founding a civil society association, etc.). In addition, professional experience outside academia, e.g. in industry, civil society or politics, was also positively highlighted in the presentation of the candidates. It was noticeable that, on the one hand, special engagement efforts from applicants were appreciated and strongly emphasized as a positive example of change agent characteristics. Yet, on the other hand, a lack of such notable engagement was not usually highlighted negatively, indicating that change agent characteristics is a rather "soft" selection criterion.

5.2 RESULTS OF THE ONLINE SURVEYS

The aim of the control group approach was to identify possible differences in change agent characteristics between funded research fellows and their bibliometric twins as these characteristics are a key evaluation and selection criterion for the funding programme. Due to the vague definition and operationalization of the concept 'change agents' in the scientific context, both in the academic literature and in the funding guidelines of the programme, the topic was approached in the two online surveys by asking about the social or developmental relevance of the research and about individual exchange, transfer and engagement activities. Furthermore, both funded and non-funded scientists were asked about their attitudes regarding priorities and motivations in their scientific work and the role science in general and scientists in particular should play in society.

The results of the comparative study are somewhat surprising. When asked "to which areas of society outside of science is your research relevant" (see Figure 1) the participants in the control group indicated a higher social relevance of their research for every social area compared to the funded research fellows (here: politics, business / industry, end-users and professionals, civil society organisations, citizens, media and culture and arts). This is surprising, since social or developmental relevance is an important assessment criterion in this research funding programme.

Similarly, in the subsequent question “Have you had any interaction with this group / these groups at any stage, before, during or after the research process?”, the participants in the control group indicated that they interacted more frequently on average with actors from all areas of society than the group of the funded research fellows did (see Figure 2).

The picture is not quite clear for the question “What kind of transfer channels do you use to make research results accessible outside of academia?” (see Figure 3). Here, the answers of the participants of the funded research fellows and the participants of the control group survey don’t differ very much with regard to the frequency of the mentioned transfer, engagement and exchange activities (research collaborations with non-scientific partners, contract research, scientific consultancy services, further education for non-scientists, commercialization activities, public engagement / citizen science, service learning, science communication), even if the frequencies indicated for these activities also tend to be somewhat higher for the control group than for the group of the funded research fellows.

On the other hand, when asked “Scientists may have different views about which goals should be given high priority in the science system. What priority do you think the following goals should have in the scientific system?”, all priorities were rated higher in the survey among the funded research fellows compared to the control group survey (see Figure 4). This includes priorities regarding responsible research, transfer and exchange activities as well as social relevance of research.

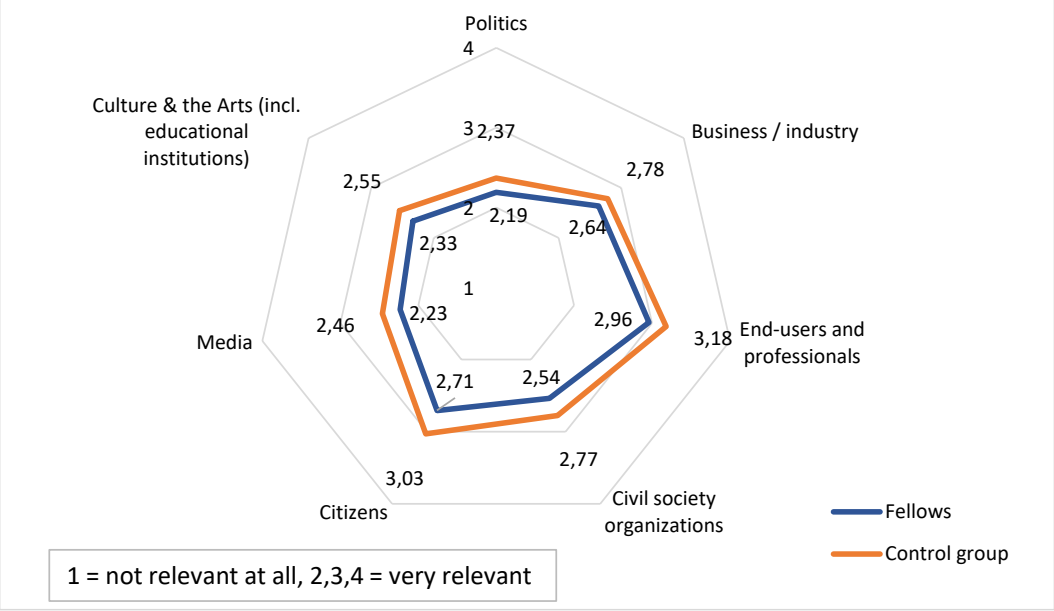

Figure 1: “To which areas of society outside of science is your research relevant?”

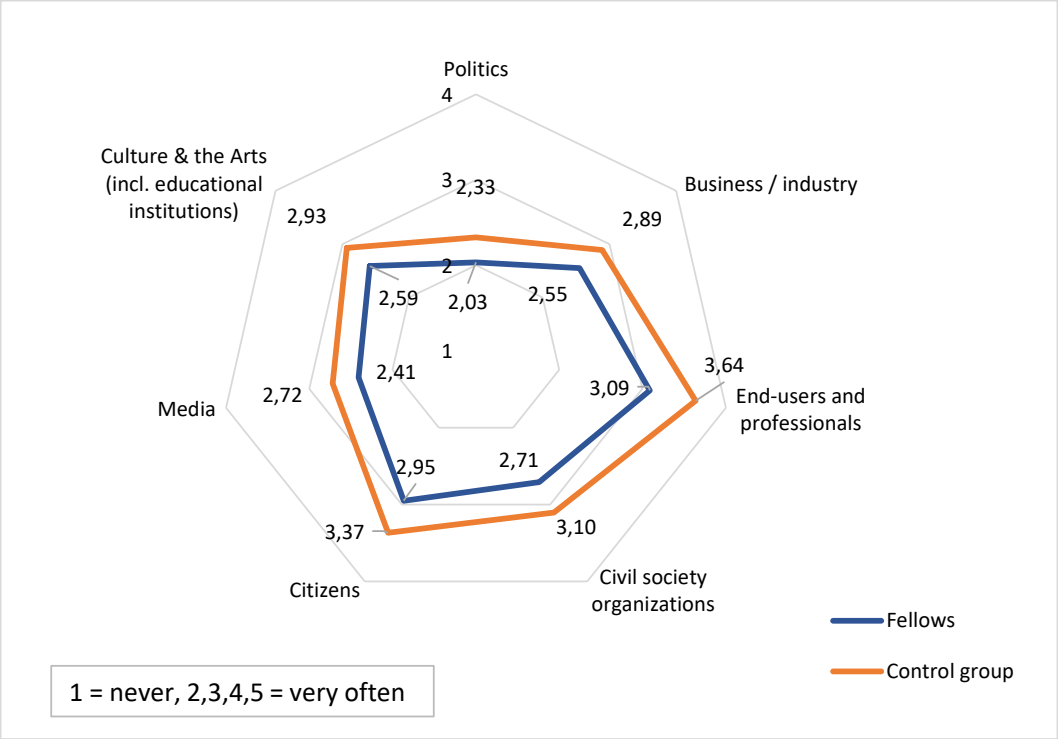

Figure 2: "Have you had any interaction with this group/these groups at any stage, before, during, or after the research process?"

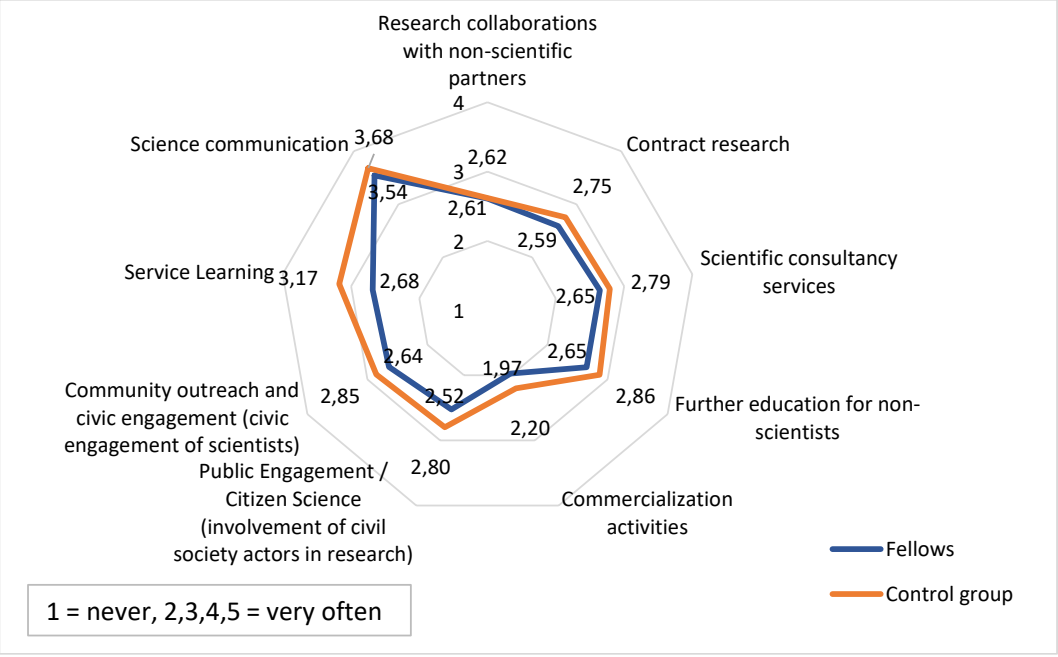

Figure 3: "What kind of transfer channels do you use to make research results accessible out-side of academia?"

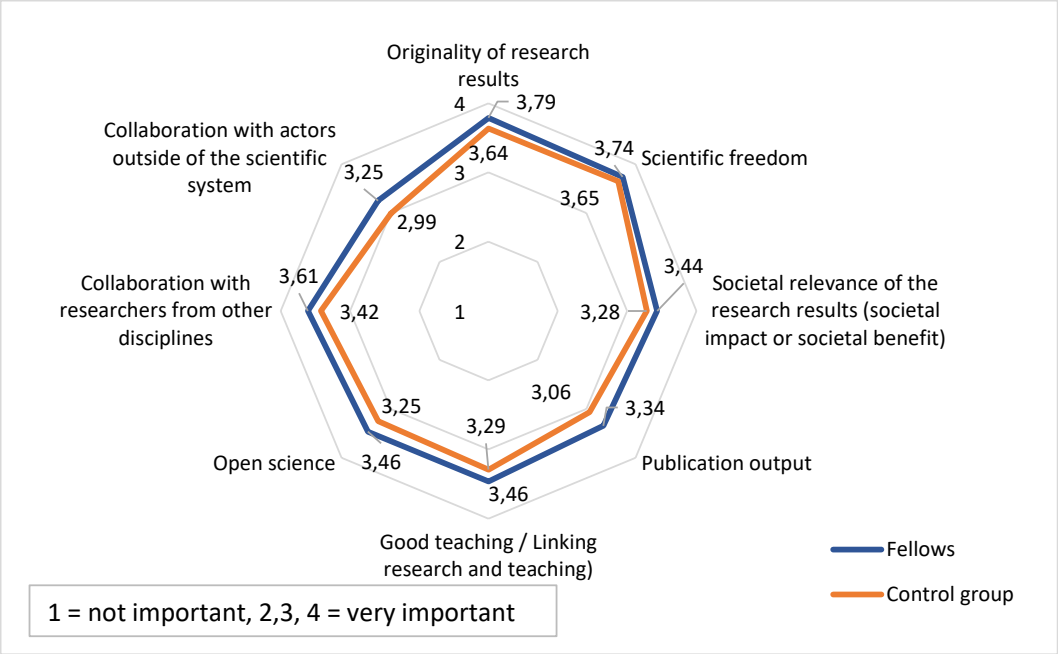

Figure 4: "What priority do you think the following goals should have in the scientific system?"

In addition, we asked scientists about their attitudes toward the role of science in society and the relationship between science and society. Here, the participants in both surveys were each asked to locate themselves between two opposing positions (semantic differential):

- Position 1a: "Science should retain a high degree of autonomy with regard to the demands of society." Position 1b: "Science should be at the service of society and provide solutions for social problems."
- Position 2a: "Scientists should actively engage in public debates." Position 2b: "Scientists should refrain from participating in public debates".
- Position 3a: "Scientists should only make statements about their own research." Position 3b: "Scientists should also contribute to topics beyond the scope of their own work."

Here too, it became obvious that the funded research fellows see science more as an important lever to promote change within society and to generate development-relevant impact. For example, the fellows were more vocal about the need for science to provide solutions to social problems, for scientists to be active in social debates and for more engagement in general. The differences are, however, statistically not significant.

Overall, the control group approach shows that the funded research fellows are more open to an active role for science in society. However, this does not manifest itself in a higher level of commitment or in a stronger engagement in transfer or exchange activities with social stakeholders.

6. DISCUSSION

This paper addresses the questions of how research funding programmes can generate both, social and scientific impact and to what extent change agent characteristics of the applying researchers can be integrated into research assessment processes. A triangulative methodological approach was used to analyze the selection process of a german research funding program targeting excellent scientists from the Global South. The funding program enables these scientists to conduct a research stay of up to two years in Germany. The effects of this research are intended to benefit the scientists' home countries. Consequently, the proposed research projects are expected to have a social and/or development relevance.

As the experience from this funding program shows, it is possible to combine funding objectives such as scientific excellence on the one hand and social impact or - as in this case - developmental relevance on the other. When implementing these funding objectives in the research assessment process, a prioritization and operationalization of the selection criteria appears necessary. In this program, development relevance or foreseeable social benefits is a necessary condition for funding and an initial selection criterion. Subsequently, the scientific excellence of the applicant and the proposed research project is evaluated. It seems particularly expedient to contextualize both selection criteria and to operationalize them as broadly as possible in order to comply with the heterogeneity of the researchers and their individual and region-specific conditions. In this program, the applicants were assessed in comprehensive form. Single criterions were rarely decisive for the assessment. Instead, all selection criteria - academic career, scientific productivity, key publications and developmental relevance - were considered with regard to the individual context. Characteristics such as age, gender, origin as well as private life circumstances were explicitly taken into account.

At the same time, the conceptualization of social and/or developmental relevance remains vague. Almost all presented and discussed applications in the research assessment process were considered to have social and/or development relevance in the proposed research projects. However, the

assessment of development relevance often appeared to be subjective. A scheme for assessing development relevance based on clear indicators appears to be necessary (e.g. relevance to which SDGs or which social problems? Relevance for which social actors or area?). Nevertheless, similar to scientific relevance, a contextualized and broad understanding of development relevance seems to be of central importance here as well. For example, to narrow of a focus on topics of classical development cooperation - in the discussion among research assessment jury members, reference was often made to typical "bridge and well projects" in poor countries - would fall short. Also, a strong focus on immediate social outputs and outcomes would have the negative effect of promoting predominantly application-oriented research. More fundamentally-oriented research or research projects from certain scientific disciplines (e.g. from humanities or natural sciences) would tend to be disadvantaged. This can contribute to a negative trend, as research in the Global South, compared to the wealthier countries in the North, is already highly application-oriented. Rather, the developmental relevance of basic research should also be appreciated in the research assessment process.

As a further step, so-called change agent characteristics of the applicants were evaluated in this funding programme. As the analyses show, clearer criteria here are also important with respect to what exactly is meant by change agent characteristics and what their significance is for the selection process. As observed in the research assessment jury sessions, change agent characteristics were mostly equated with special scientific or (civic) social engagement. It was positively emphasized in several applications. However, the lack of such an engagement was not evaluated negatively. However, the surveys showed that change agent characteristics are a possible approach to increase the likelihood of realizing social impact through research funding and to connect scientific excellence with developmental and/or social relevance. According to the control group survey, the funded scientists had significantly stronger change agent attitudes than the non-funded researchers with regard to the question of what role science should play in society. This refers in particular to the statement (i) science should be at the service of society and provide solutions for societal problems, (ii) scientists should actively engage in public debates and (iii) scientists should contribute to topics beyond the scope of their own work. These positions received higher approval ratings among the funded research fellows than among the non-funded control group. This indicates that the funding programme is quite successful in selecting those scientists who support an image of science that is actively involved in society and shapes social developments. However, this did not manifest itself in higher engagement or transfer activities, which were used as indicators

for change agency in this study. Nevertheless, it is important in this context to point out the limitations of the study design, especially the methodological approach to generating the comparison group using bibliometric twins. It was not possible to form a fully representative comparison group, as for some of the research fellows no identical bibliometric twin could be identified. This also shows the possibilities and limitations of control group approaches in many real evaluation contexts, which are increasingly demanded by the evaluation community when analysing science, technology and innovation funding measures (EFI 2024). Due to the individual design of the programme and the numerous influencing factors and characteristics of the treatment group that need to be considered, it is difficult to identify a representative control group - at least in this case.

At the same time, the application of selection criteria in research funding processes that go beyond purely scientific aspects of the application or the scientific performance of the researchers and consider personal characteristics or engagement activities should be handled with great caution. They can also harbour risks, especially if they are not clearly conceptualized and defined. For example, negative incentives could arise for scientists to overload their research proposals with unfounded and completely exaggerated expectations regarding societal impact. Reviewers could be distracted from other selection criteria such as originality or scientific standards by too much story-telling, which could result in more scientists being selected who mainly master the right proposal prose. In addition, if there is too much focus on aspects like development potential or change agent criteria, there could be a risk that in the end, mostly strongly application-orientated research is funded to the detriment of basic research, which is primarily interest-orientated rather than application-orientated., as was also critically noted in the discussion among the jury members. Future research should therefore focus more on the analysis of change agent characteristics among scientists, as there is a considerable need of empirical and conceptual work in this area.

REFERENCES

Alberts, B.; Kirschner, M. W.; Tilghman, S.; Varmus, H. (2014) Rescuing US biomedical research from its systemic flaws. In *Proceedings of the National Academy of Sciences* 111/16: 5773–7.

Beckhard, R. (1969) *Organizational development: strategies and methods*. Boston: Addison - Wesley.

Belcher, B. M.; Rasmussen, K. E.; Kemshaw, M. R. et al. (2016) Defining and assessing research quality in transdisciplinary context, *Research Evaluation* 45/1: 1–17.

Bonn, N. A.; Bouter, L. (2023) Research Assessments Should Recognize Responsible Research Practices. Narrative Review of a Lively Debate and Promising Developments. In: Erick Valdés, Juan Alberto Lecaros (eds.) *Handbook of Bioethical Decision. Volume I: Decision at the Bench*, pp. 441–72, Cham: Springer International Publishing.

Boon, W.; Edler, J. (2018) Demand, Challenges, and Innovation. Making Sense of New Trends in Innovation Policy*, Science and Public Policy* 45/4: 435–47.

Bornmann, L. (2013) What is social impact of research and how can it be assessed? A literature survey. In *Journal of the American Society for Information Science and Technology* 64/2: 217–33.

Caldwell, R. (2005) Things fall apart? Discourses on agency and change in organizations. In *Human Relations* 58/1: 83–114.

Curry, S.; Rijcke, S. de; Hatch, A.; et al. (2020) The changing role of funders in responsible research assessment: progress, obstacles and the way ahead. Working Paper. Research on Research Institute (RoRI).

D'Este, P.; Ramos-Vielba, I.; Woolley, R.; Amara, N. (2018) How do researchers generate scientific and societal impacts? Toward an analytical and operational framework, *Science and Public Policy* 45/6: 752–63.

Donovan, C.; Hanney, S. (2011) The 'Payback Framework' Explained, *Research Evaluation* 20: 181–3.

DORA (2021) *Reimaging academic assessment: Stories of innovation and change*. Available online at https://sfdora.org/dora-case- studies/.

Edwards, M. A.; Roy, S. (2017) Academic Research in the 21st Century: Maintaining Scientific Integrity in a Climate of perverse Incentives and Hypercompetition, *Environmental Engineering Science* 31/1: 51–61.

EFI - Commission of Experts for Research and Innovation (2024): Report on Research, Innovation and Technological Performance in Germany 2024, Berlin: EFI.

Ferretti, F.; Guimaraes Pereira, D. V.; Hardemann, S. (2018) Research excellence indicators: time to reimage the 'making of', *Science and Public Policy* 45/5: 731–41.

Gingras, Yves (2014) *Bibliometrics and Research Evaluation. Uses and abuses.* Cambridge: MIT Press.

Hagstrom, Warren O. (1975) *The Scientific Community*. Carbondale: Southern Illinois University Press.

Hicks, D.; Wouters, P.; Waltman, L.; Rijcke, S. d.; Rafols, I. (2015) The Leiden manifesto for research metrics, *Nature* 520, 429–31.

Hutton, David W. (1994) *The Change Agents' Handbook*. Milwaukee: ASQC/Quality Press.

Joly, P. B.; Gaunand, A.; Colinet, L.; Larédo, P.; et al. (2015) ASIRPA: A comprehensive theory-based approach to assessing the societal impacts of a research organization, *Research Evaluation* 24/4: 440–53.

Kawulich, B. B. (2005) Participant Observation as a Data Collection Method, *Forum Qualitative Sozialforschung: Forum Qualitative Social Research* 6/2.

Kok, M. O.; Schuit, A. J. (2012) Contribution mapping: A method for mapping the contribution of research to enhance its impact, *Health research policy and systems* 10/1: 1–16.

Kraemer-Mbula, Erika; Tijssen, Robert; Wallace, Matthew L.; McLean, Robert (2020) *Transforming Research Excellence. New Ideas from the Global South*. Cape Town: African Minds.

Lauronen, J.-P. (2020) The dilemmas and uncertainties in assessing the societal impact of research, *Science and Public Policy* 47/2.

Lunenburg, F. C. (2010) Managing Change. The Role of the Change Agent, *International Journal of Management, Business, and Administration* 13/1: 1–6.

Matt, M.; Gaunand, A.; Joly, P.-B.; Colinet, L. (2017) Opening the Black Box of Impact - Ideal-Type Impact Pathways in a Public Agricultural Research Organization, *Research Policy* 46: 207–18.

Mazzucato, M. (2018) Mission-oriented innovation policies: challenges and opportunities, *Industrial and corporate change* 27/5: 803–15.

Merton, R. K. (1957) Priorities in scientific discovery: A chapter in the sociology of science, *American Sociology Review* 22: 635–59.

Moed, H. F. (2020) Appropriate Use of Metrics in Research Assessment of Autonomous Academic Institutions, *Scholarly Assessment Reports* 2/1.

Moher, D.; Naudet, F.; Cristea, I. A.; Miedema, F.; et al. (2018) Assessing scientists for hiring, promotion, and tenure, *PLoS Biology* 16/3.

Moran, J.; Brightman, B. (2000) Leading organizational change, *Journal of Workplace Learning* 12: 66–74.

Ottaway, R. N. (1983) The Change Agent: A Taxanomy in Relation to the Change Process, *Human Relations* 36/4: 361–92.

Reale, E.; Avramov, D.; Canhial, K.; Donovan, C.; et al. (2017) A review of literature on evaluating the scientific, social and political impact of social sciences and humanities research, *Research Evaluation* 27/4: 298–308.

Rogers, E. M.; Shoemaker, F. F. (1971) *Communications of innovations: A cross-cultural appraoch*. New York: The Free Press.

Smit, J. P.; Hessels, L. K. (2021) The production of scientific and societal value in research evaluation: A review of societal impact assessment methods, *Research Evaluation* 30/3: 323–35.

Spaapen, J.; van Drooge, L. (2011) Introducing 'Productive Interactions' in Social Impact Assessment, *Research Assessment* 20/3: 211–18.

Tijssen, R.; Winnink, J. (2022) Global and Local Research Excellence in Afrika: New Performance Assessment and Funding, *Science, Technology and Society* 27/3: 368–87.

UN General Assembly (2015) *Transforming our world. The 2030 Agenda for Sustainable Development*. 21 October 2015. A/RES/70/1. New York.

Wilsdon, J.; Allen, L.; Belfiore, E.; Campbell, P.; et al. (2015) *The metric tide: Report of the independent review of the role of metrics in research assessment*

and management. Available online at https://re.ukri.org/documents/hefce-documents/metric-tide- 2015- pdf/.

Zuckerman, H.; Merton, R. K. (1971) Patterns of evaluation in science: Institutionalisation, structure and functions of the referee system, *Minerva* 9/1: 66–100.

AUTHOR

HENDRIK BERGHÄUSER
Fraunhofer Institute for Systems and Innovation Research ISI
Breslauer Straße 48
D-76139 Karlsruhe
E-Mail: Hendrik.Berghaeuser@isi.fraunhofer.de
ORCID: 0000-0002-3884-3404