

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-

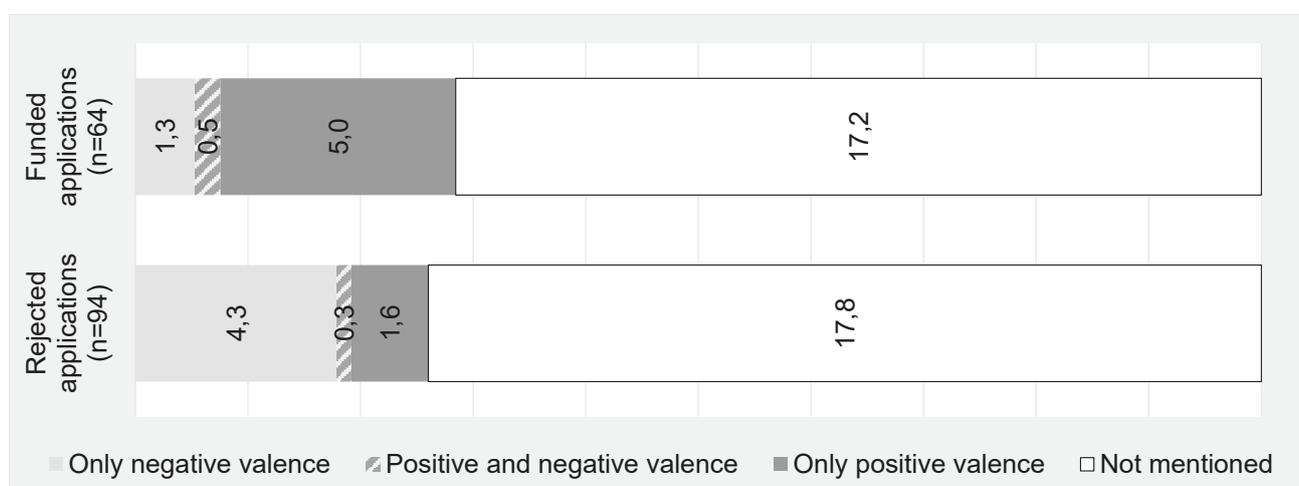


Figure 1: Average number and valence of rated dimensions (total 24)

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.

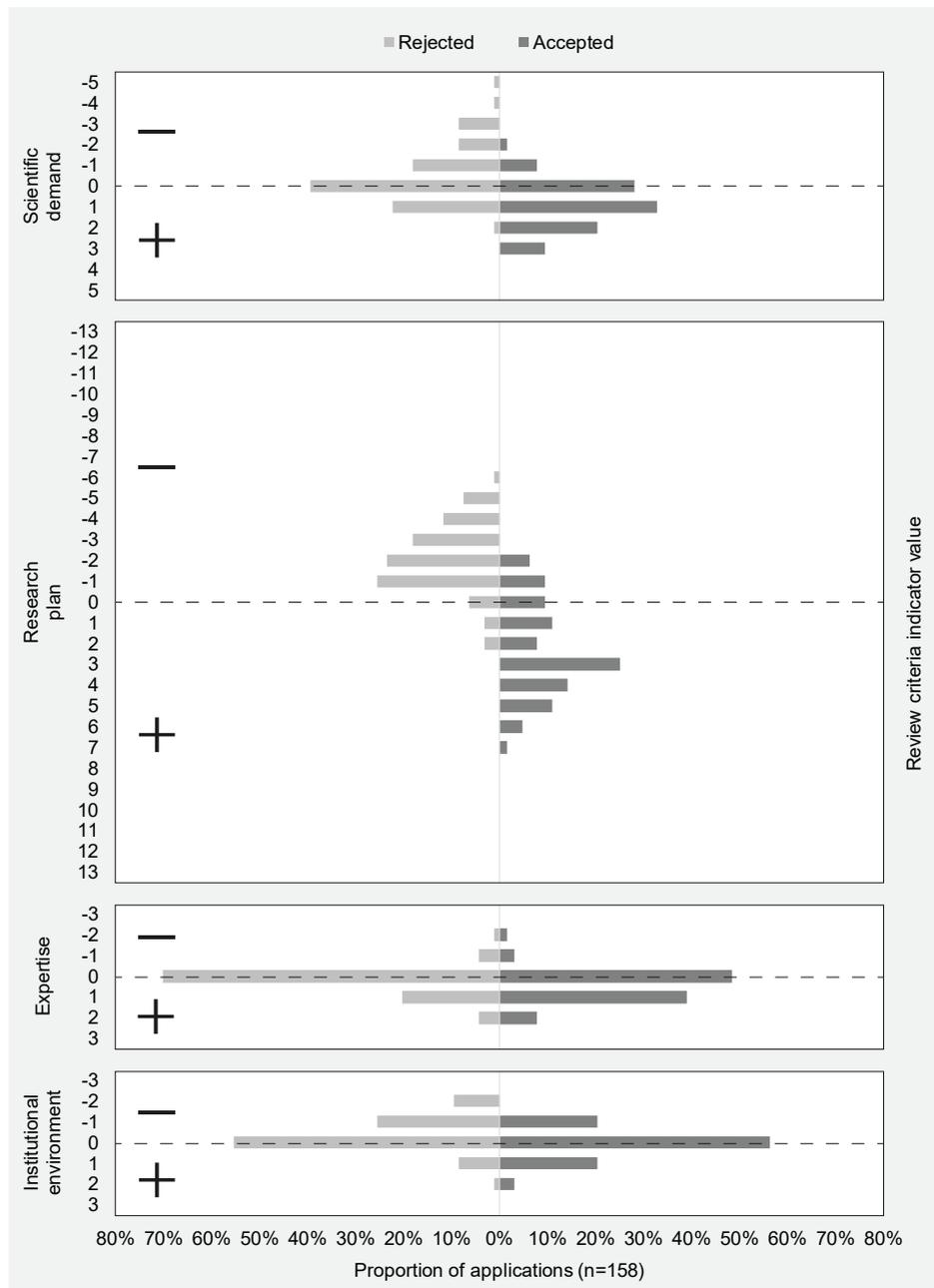


Figure 2: Distribution of the four review criteria indicators by funding decision

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).

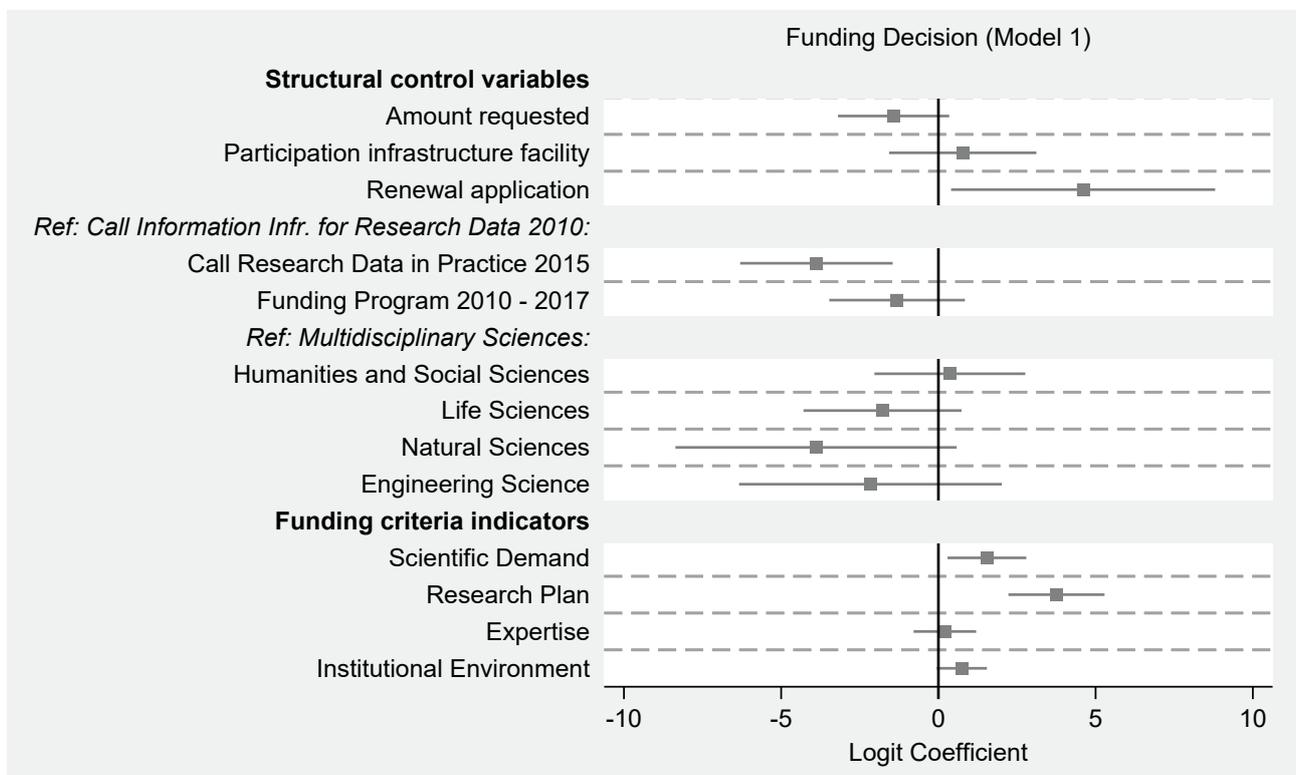


Figure 3: Logit coefficient plot model 1 funding decision (compare table 5)

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