

Exploring the evaluation of inter- and transdisciplinary research proposals: Lessons from Dutch research funding reform

Jochem Zuiderwijk^{*}, Wolfgang Kaltenbrunner^{**} and Lotte Krabbenborg^{***}

^{*}*jochem.zuiderwijk@ru.nl*

ORCID (0000-0002-0896-1063)

Institute for Science in Society (ISiS), Radboud University, Netherlands

^{**}*w.kaltenbrunner@cwts.leidenuniv.nl*

ORCID (0000-0003-4311-8212)

Centre for Science and Technology Studies (CWTS), Leiden University, Netherlands

^{***}*lotte.krabbenborg@ru.nl*

ORCID tbd

Institute for Science in Society (ISiS), Radboud University, Netherlands

Abstract: This paper presents part of the preliminary results of a pilot research project commissioned by the Dutch research funder NWO, aimed at studying the current assessment of interdisciplinary research proposals at NWO and make possible recommendations for improvement based on the international literature. In the presented study a combination of document analysis, semi-structured interviewing and observation was used to study current assessment practices in the context of three different funding calls to which inter- or trans-disciplinary research proposals were submitted. Findings from the three calls were compared to best practices presented in the existing literature and used to make recommendations for possible improvements, reflect on possible barriers for effective reforms, and suggest further research avenues for a larger potential follow-up project.

1. Introduction

Interdisciplinary research has become increasingly common, whilst policy makers and research funding bodies have increasingly embraced the potential of inter- and transdisciplinary research (ITDR) to tackle larger societal challenges (OECD, 2020; Schneider et al., 2019). In this context the Dutch national research funder NWO has developed a variety of calls explicitly aimed at tackling societal challenges and complex problems through ITDR, whilst interdisciplinary research proposals are also increasingly submitted to its open funding calls. However, the international literature shows there is still much unclear about how best to evaluate ITDR and there are broader worries about the efficacy of current assessment frameworks to adequately deal with ITDR proposals (e.g. Laursen et al., 2022; McLeish & Strang, 2016). Within this context NWO's Knowledge Platform for Inter- and Transdisciplinary Research¹ has set up a pilot research project aimed at investigating the current state and potential need for modifications of the assessment of ITDR proposals at NWO, and compiling a set of relevant recommendations available from both academic literature and international experience. This paper presents the results of one sub-project in this pilot, which was aimed at the qualitative exploration of current ITDR assessment practice. A variety of qualitative research methods was used to take stock of current evaluation practices around ITDR in a handful of funding calls, identify any particular issues and best practices, and relate findings to suggestions for ITDR evaluation found in the academic literature in order to make a preliminary set of general recommendations. This paper seeks to discuss the preliminary results of this exploratory study and reflects on the broader relevance and potential lessons that can be learned from the results at this stage of the project.

¹ <https://www.nwo.nl/en/knowledge-platform-for-interdisciplinary-and-transdisciplinary-research>

Note: Due to an embargo period placed on the data and results of this study until they have been published in an internal report at the end of June, our final analysis and recommendations have been removed from this paper until the final presentation.

2. Methods and case study selection

2.1. Methods and data

The goal of this sub-project was to employ qualitative methods to explore the practices through which ITDR proposals are assessed in various different calls at NWO, take first stock of both potential issues and effective practices, and compare and relate these to the research and best practices found in the relevant academic literature on ITDR assessment. For this purpose the research team made a selection of three funding calls and conducted:

- Document analysis of all associated reviewer instructions, review reports and proposals
- 11 semi-structured interviews with NWO staff and assessment committee members
- 6 observations at committee preparation and final assessment meetings for all three calls.

Document analysis was used to understand the various call procedures, create an overview of all the information and data available to reviewers, and analyse reviewer comments on proposals. Interviews focused on discussing each stage of the assessment procedure, any potential issues experienced by the interviewees, and reflecting on the observed meetings. All interviews were recorded, transcribed and coded through a simple coding scheme based on initial axial coding and themes gathered from the literature. Observations were made by three different researchers using a concise observation protocol.

2.2. Case studies

To explore a variety of evaluation procedures the qualitative research project included three different call procedures from three different funding programs that were finishing within the pilot project runtime. A selection of three calls was made to cover a variety of funding programs, goals and procedures that deal with ITDR proposals at NWO. The three selected calls were each part of a different kind of funding program. Whereas call #1 was part of a continuous open call, call #2 was part of a set of mission-oriented calls, and call #3 was a thematic call derived from the Dutch Research Agenda (NWA). Whereas the first type of call is typically meant for more ‘bottom-up’ research proposals by individuals or groups, the other two tend to involve calls driven by societal questions or issues that typically require larger inter- or transdisciplinary research consortia.

Table 1. Steps in the evaluation procedures of the three studied calls

Phase	Call #1	Call #2	Call #3
Preliminary stage	None	Pre-proposal submitted	Initiative submitted
		Pre-proposal assessed in committee meeting	Initiative shared publicly
		Matchmaking workshop for applicants	Matchmaking workshop for applicants
Assessment stage	External review reports		
	Written rebuttal by applicant		
	Committee instructed via email	Committee instructed via introduction meeting	Committee instructed via introduction meeting
	Cluster committee meeting	Committee meeting with applicant interview	Committee meeting with applicant interview
	Domain committee meeting		
	Final scoring	Final scoring	Final scoring

Each of the researched calls involved a different assessment procedure (see Table 1). Within established parameters, calls are tailored to particular goals and questions set for the individual calls. This means that the exact nature of the procedures described here are not necessarily shared by all calls within their programs, and there is in fact quite a bit of differentiation to be found between procedures.

3. Results

3.1. Funding call #1

Funding call #1 concerned a continuous open call for basic research proposals from exact science fields. Although the call has generally been set up to assess disciplinary proposals, it explicitly allows for interdisciplinary proposals to be submitted as well. Funding agency staff seek out two to three external reviewers per proposal to write out individual referee reports without scoring. Applicants are then given a right to reply to the reports in written form. Proposals, together with their reports and rebuttal, are then grouped into clusters of topical and disciplinary affinity and further assessed by a cluster committee in which reviewers from the relevant disciplines are grouped together. Proposals are assigned to two or three committee members who provide a 'pre-advice' on each proposal prior to meeting, based on a set of written instructions, and give a numerical scoring for each given criterium. The committee subsequently discusses each proposal for around three to five minutes in a collective meeting, in which discussion and final scoring can generally only be based on arguments already given in the referee reports and rebuttals. The emphasis in this assessment procedure is placed on the assessment of the external reviewers in the first phase, while the role of the committee assessment in the second phase is closer to that of a panel judging and scoring the reports. The final scoring and subsequent rankings of the different cluster committees are subsequently combined for review by an overarching exact sciences committee that decides on the overall final ranking and funding allocation.

For every stage of the process the organizational staff managing the call procedure is always tasked with seeking out reviewers with the relevant expertise to assess proposals. In principle, this also means they have to look for relevant interdisciplinary experts to assess interdisciplinary proposals. In practice, finding these particular interdisciplinary experts for the referee report stage can prove difficult, whereas the mixed nature of the subsequent cluster committees means staff generally have to focus on trying to 'cover' each discipline presented in the multiple proposals to be assessed collectively at this stage, by adding and using a variety of individual experts from each area. Staff consider interdisciplinary proposals submitted to this call to be rare. If they *are* submitted, it is quite likely they will be mostly assessed by (mono-)disciplinary experts from multiple fields in both stages of the process. Because the call is generally set up for basic research from distinct disciplines there is also no explicit consideration or criterium focused on interdisciplinary collaboration included in the procedure, although discussion of proposals with more than one applicant includes a brief extra consideration of the proposed setup of their 'collaboration'.

3.2. Funding calls #2 and #3

Funding calls #2 and #3 both entailed calls focused on complex societal issues which require inter- and transdisciplinary research consortia to be addressed adequately and their respective assessment procedures showed significant similarities. When compared to call #1 they both entailed a longer run up phase to the full proposal submission and evaluation, in which different forms of pre-proposals were submitted and matchmaking events were used to assist applicants in forming the larger consortia asked for in the call. For call #2 applicants submitted a pre-

proposal that was given a basic go/no-go advice to develop a full proposal, which was given on the basis of an assessment by the same committee as was tasked with the final assessment of the full proposals. For call #3 a form of pre-proposal submission came in the form of submitting an 'initiative' description that was not assessed (beyond a check of basic requirements), but rather published publicly to also assist in finding further partners for consortia formation.

These two calls did not make use of external reviewer reports, and the right to hearing and rebuttal was fulfilled through committee interviews with applicants rather than written reply. This means the committee assessment was responsible for the core assessment and scoring of the proposals in these procedures. Similar to call #1 proposals were assigned to committee members to provide pre-advice and initial scoring, with space in the final meeting to re-score proposals on the basis of committee discussion and the interviews with applicants. Aside from written instructions all committee members attended an introductory 'calibration' meeting in which the assessment procedure and criteria were discussed in both calls. In call #2 this calibration took place after pre-proposals had already been assessed for a recommendation. In call #3 a very tight time schedule meant this calibration took place very shortly before the pre-advice and initial scoring had to be submitted and the final meeting was coming up, whilst reviewers had to have already read the proposals to submit interview questions.

Because both calls explicitly asked for research consortia involving multiple different disciplinary and non-academic partners, the assessment committee also consisted of a variety of academic and non-academic reviewers, as is dictated by the funders policy. Criteria and sub-criteria in both calls entailed assessing various aspects of interdisciplinary and transdisciplinary collaboration such as the feasibility of collaborative work, inter- and transdisciplinary plans for societal and scientific impact, the match between call goals and consortium composition, and the inclusion and role of different types of disciplines and actors. More detailed questions on how the different disciplines and forms of knowledge would be integrated seemed to be more difficult for committee members to assess based on the given proposals and interviews, and interview questions asked by committee members tended to be mostly focused on specialist issues related to their own disciplinary areas rather than the overall consortia. In relation to call #1 much fewer proposals were discussed, which allowed for much more time for in depth discussion of each criterium.

4. Discussion

The academic literature concerning the evaluation of inter- and transdisciplinary research has come to present a broad range of recommendations and best practices of varying levels of detail (e.g. Lyall & King, 2013; McLeish & Strang, 2016; Pohl et al., 2011). On the basis of the most pertinent issues and questions arising from the collected data concerning the evaluation procedures used to assess ITDR at NWO, a range of relevant recommendations across these literary sources was selected, compared and distilled into five general recommendations.

These recommendations represent five general lessons and talking points that could prove useful for any funding institution looking to adapt their programs and policies to accommodate interdisciplinary and transdisciplinary research. The fifth and final recommendation pertains specifically to the evaluation of research in which non-academic actors, such as private partners or public organizations, are involved (transdisciplinary research).

4.1. The evaluation of interdisciplinary research requires a distinct approach

Many overviews of the literature indicate that the evaluation of interdisciplinary research needs to be structured in a way that the particular added value of interdisciplinary research as a

potential new ‘transformative and emergent whole’ can be recognized (e.g. McLeish & Strang, 2016, p. 6). Existing forms of peer review are generally found to be ill-suited for this purpose and require significant adjustments (e.g. European Union Research Advisory Board, 2004; GRC, 2016; Institute of Medicine, 2004; Mayer et al., 2013; Vienni Baptista et al., 2020). Simply adding an extra layer of ‘interdisciplinary evaluation’ to existing procedures of disciplinary evaluation has been found to be insufficient in this context (e.g. McLeish & Strang, 2016, p. 4). Although such added layers might lead to more recognition of ITDR’s value, it generally turns multiple dimensions of interdisciplinary work into an *extra* criterium or demand *on top* of multiple accumulated sets of critical monodisciplinary evaluative frames to which proposals are also subjected. The evaluation of ITDR has thus been shown to do better in calls specifically meant for ITDR proposals (Lyall & King, 2013, p. 13). This points to an issue for open calls in which ITDR proposals are included within evaluation procedures that are generally set up for proposals from more traditional disciplines, such as with call #1.

4.2. Assessment of ITDR proposals requires extra reviewer preparation and calibration

For the adequate assessment of ITDR proposals additional preparation for reviewers and assessment committees is regularly presented as a crucial intervention. The relevant literature holds that extra training, instruction and interaction is required for reviewers to ‘calibrate’ and come to a shared understanding of the evaluation process, the assessment criteria, the desired form of interdisciplinary collaboration in proposals and an interdisciplinary notion of research quality (e.g. Burgess et al., 2016, p. 4; Lyall et al., 2013, p. 68; McLeish & Strang, 2016, p. 5; Ridley, 2016, p. 8). It has been suggested that this requires at least one extra preparatory meeting prior to the actual assessment (e.g. Lyall & King, 2013: 13). These suggestions are difficult to implement in multiple stage procedures where individual review reports have a central role, as is the case with call #1. Calls #2 and #3 both involved preparatory meetings in line with these recommendations, although the time planning of the call procedures meant these preparatory meetings were held while part of the assessment was already underway, which arguably hampers their effectiveness.

4.3. Combining multiple disciplinary experts does not lead to good ITDR evaluation

Research shows that combining experts from the multiple disciplines presented in an ITDR proposal within an assessment committee does not in itself lead to the effective assessment of the new whole presented in these proposals (Bruun et al., 2005; McLeish & Strang, 2016, p. 4). A setup with multiple ‘monodisciplinary’ experts like this can become extra problematic when assessment is spread over multiple stages that do not all include space for deliberation between reviewers, which together creates the risk of ‘double, or multiple, jeopardy in the sequential evaluation of IDR through single-disciplinary lenses’ (McLeish & Strang, 2016, p. 3). In this context it has been noted that ‘[w]hile single-discipline experts have an important place in IDR evaluation, their role ought to be supportive of those chosen for their ability to judge the critical “emergent” outcomes of the research. This is the single intervention most commonly reported as effective’ (ibid., p. 6). A setup like call #1 that emphasizes individual reviewer reports tends to make the type of interaction this requires difficult to implement. For procedures such as those presented in calls #2 and #3 this suggestion re-emphasizes the need to include enough interdisciplinary experts within review committees.

4.4 Pay explicit attention to interdisciplinarity and integration in interviews and discussions

The type and degree of integration between the multiple disciplines and types of knowledge involved in a proposal, and thus of the researchers and stakeholders involved, are both central themes in the effective assessment of ITDR proposals (e.g. Laursen et al., 2022). To keep focus on these elements it has been suggested that review committees should always consist of a mix

of ‘specialists and generalists’ (Pohl et al., 2011, p. 8) and should always include enough reviewers with knowledge and experience of interdisciplinary research (Lyall et al., 2011). To prevent situations where the inclusion of interdisciplinary experts amounts to a purely symbolic measure, it is important to give one or more of these committee members the explicit task of keeping focus on criteria of interdisciplinary setup and integration throughout all phases of the assessment procedure, and make sure these aspects of the assessment are not discarded in favor of more traditional disciplinary criteria (Lyall & King, 2013, pp. 13–14). For calls tailored to ITDR research, such as calls #2 and #3, it might prove useful to provide a separate or additional training for reviewers tasked explicitly with assessing particular ITDR elements, such as the integration of disciplines and forms of expertise.

4.5 Address existing barriers for participation in research assessment committees

The makeup of review committees might be the most important factor in effectively assessing ITDR research (Lyall et al., 2011; McLeish & Strang, 2016, p. 5). When it comes to transdisciplinary research this means it is crucial to include diverse forms of knowledge, experience and perspectives of relevant societal actors outside of academia when forming committees as well, as is already the policy in some of the call procedures discussed in the results above. However, there can be various practical barriers complicating the inclusion of particular societal parties in committees, many of which include a disparity in terms of the resources, time or independence of different kinds of organizations and individuals. For calls #2 and #3 for example, difficulties were mentioned with finding willing and able representatives of particular public and private organizations due to the very limited time availability of public officials or the relatively limited nature of the available stipends in comparison with remunerations provided for work in the private sector. These barriers might not only lead to gaps in the perspectives included in review committees, but can also lead to risks of so-called ‘elite capture’ by larger or more wealthy parties gaining a disproportionate amount of control over the direction or flow of research funding (e.g. Bender, 2022). Moreover, if we also consider insights from work on transdisciplinary decision making in general, we know that even when more practical barriers for participation are overcome, the actual role and input of certain actors may subsequently also be impacted through different relations of power or epistemological and ontological positions (e.g. Ludwig & Boogaard, 2021, p. 26). In other words, effective committee assessment of ITDR proposals requires serious consideration of both the makeup of committees and the practical role subsequently given to the disparate members involved once they participate.

4.6. Barriers for reform

The research also encountered a variety of elements that could form potential barriers to efforts of improving the assessment of interdisciplinary research at NWO. An example of this is the large amount of information that reviewers are already given in preparation of assessment, which hampers the effectiveness of any interventions that would simply include adding more instructions. Besides the general instructions and explanation of the assessment and the criteria, committee members were given information on a host of other issues including guidelines on aspects of ‘diversity and inclusion’ in the assessment process, the rules set out by the DORA declaration, and instructions for using the funder’s software systems. The current degree of information overload amongst reviewers, which could risk the effectiveness and reliability of the evaluation process, is a shared concern that has to be taken into account in interventions. To be effective these interventions would have to focus on communicating information differently, rather than merely add to reviewer instructions. Further discussion of potential barriers will be added to the final version of this paper with the release of the embargo.

A second element concerns the technological infrastructure of the funder. In particular, the software system used in part of all submission processes relies on more ‘traditional’ disciplinary codes in ways that obstruct a clear view on the real (inter)disciplinary nature of proposals and although efforts are underway to change or replace it with a more up-to-date and adaptable system, this has proven to be a slow and difficult process (see *technomaterial resilience*). This complicates suggested efforts to guide applicants towards a suitable programme with a procedure aligned with a proposal’s particular mono-, inter- or transdisciplinary nature.

5. Conclusion

The research team responsible for this study was tasked with conducting an exploratory qualitative study of current practices around the assessment of ITDR proposals at the Dutch funder NWO, and see if and how particular improvements might be made based on the relevant international literature. In the final presentation of this paper, after the embargo period, we will provide our comparison of current practices and best practices in the literature, and present our final conclusions and recommendations.

In the context of NWO's current assessment of ITDR proposals, five general recommendations were compiled, based on best practices mentioned in the literature, and these were subsequently used to reflect on current assessment practices at NWO. Whereas the two calls aimed at particular societal issues and questions already adopted numerous practices recommended in the literature, the open call was essentially tailored to the assessment of the majority of proposals submitted in the call, which are monodisciplinary. This makes it more difficult to suggest improvements through the implementation of particular best practices and leaves the question whether such open calls should be overhauled or new open calls tailored to ITDR proposals should be considered.

Open science practices

Due to the formal confidential nature of both the assessed research proposals and the assessment procedures examined in this pilot project, the data used in this analysis has not been made openly available to the public. The NWO has also restricted the publication of the results and recommendations until both have been published internally

Acknowledgments

In the development of this research project we owe our thanks to our colleagues Laurens Hessels, Daniël Lakens, Anne-Floor Scholvinck, and Krist Vaesen, who worked with us on the NWO's overall research pilot project. Moreover, the existence of this research project is only possible thanks to NWO, the team behind NWO's Knowledge Platform for Inter- and Transdisciplinary Research, in particular Aniek van den Eersten, and the various staff and committee members that have chosen to participate in our research.

Author contributions (following CrediT taxonomy)

Jochem Zuiderwijk: Conceptualization, Methodology, Investigation, Data Curation, Writing – Original Draft, Writing – Review & Editing, Project Administration

Wolfgang Kaltenbrunner: Conceptualization, Methodology, Investigation

Lotte Krabbenborg: Conceptualization, Methodology, Investigation

Competing interests & funding information

Both the author and the broader research team responsible for the pilot project were seconded to the NWO for this project and the project as a whole was funded by NWO itself.

References

- Bender, K. (2022). Research–Practice–Collaborations in International Sustainable Development and Knowledge Production: Reflections from a Political-Economic Perspective. *The European Journal of Development Research*, 34(4), 1691–1703.
- Bruun, H., Hukkinen, J. I., Huutoniemi, K. I., & Thompson Klein, J. (2005). Promoting interdisciplinary research: The case of the Academy of Finland. *Publications of the Academy of Finland, Vuosikerta. 2005*(Nro 8). <http://hdl.handle.net/10138/232267>
- Burgess, M., Garnett, V., O'Connor, A., & Ohlmeyer, J. (2016). *Interdisciplinarity for Impact* [Workshop report]. Trinity College Dublin.
- European Union Research Advisory Board. (2004). *Interdisciplinarity in Research—EURAB 04.009-FINAL* [Report]. EURAB.
- GRC. (2016). *Interdisciplinarity: Survey Report for the Global Research Council 2016 Annual Meeting* [Survey report]. Global Research Council - DJS Research. https://www.globalresearchcouncil.org/fileadmin/documents/GRC_Publications/Interdisciplinarity_Report_for_GRC_DJS_Research.pdf
- Institute of Medicine. (2004). *Facilitating Interdisciplinary Research*. National Academies Press. <https://doi.org/10.17226/11153>
- Laursen, B. K., Motzer, N., & Anderson, K. J. (2022). Pathways for assessing interdisciplinarity: A systematic review. *Research Evaluation*, 31(3), 326–343.
- Ludwig, D., & Boogaard, B. K. (2021). Making transdisciplinarity work: An epistemology of inclusive development and innovation. In *The Politics of Knowledge in Inclusive Development and Innovation* (pp. 19–33). Routledge.
- Lyall, C., Bruce, A., Marsden, W., & Meagher, L. (2013). The role of funding agencies in creating interdisciplinary knowledge. *Science and Public Policy*, 40(1), 62–71. <https://doi.org/10.1093/scipol/scs121>
- Lyall, C., & King, E. (2013). *International good practice in the peer review of interdisciplinary research* [Commissioned report]. University of Edinburgh. <https://www.research.ed.ac.uk/en/publications/international-good-practice-in-the-peer-review-of-interdisciplina>
- Lyall, C., Tait, J., Meagher, L., Bruce, A., & Marsden, W. (2011). *A Short Guide to Evaluating Interdisciplinary Research*. <https://www.research.ed.ac.uk/en/publications/a-short-guide-to-evaluating-interdisciplinary-research>
- Mayer, K., König, T., & Tuménas, R. (2013). *Horizons for Social Sciences and Humanities: Conference Report* [Conference Report]. Mykolas Romeris University Publishing.
- McLeish, T., & Strang, V. (2016). Evaluating interdisciplinary research: The elephant in the peer-reviewers' room. *Palgrave Communications*, 2(1), Article 1. <https://doi.org/10.1057/palcomms.2016.55>

OECD. (2020). *Addressing societal challenges using transdisciplinary research* (OECD Science, Technology and Industry Policy Papers No. 88; OECD Science, Technology and Industry Policy Papers, Vol. 88). <https://doi.org/10.1787/0ca0ca45-en>

Pohl, C., Perrig-Chiello, P., Butz, B., Hadorn, G. H., Joye, D., Lawrence, R., Nentwich, M., Paulsen, T., Rossini, M., & Truffer, B. (2011). *Questions to evaluate inter-and transdisciplinary research proposals* [Working paper]. td-net for Transdisciplinary Research.

Ridley, A. (2016). *Improving recognition of team science contributions in biomedical research careers* [Report]. Academy of Medical Sciences.
<https://acmedsci.ac.uk/policy/policy-projects/team-science>

Schneider, F., Buser, T., Keller, R., Tribaldos, T., & Rist, S. (2019). Research funding programmes aiming for societal transformations: Ten key stages. *Science and Public Policy*, 46(3), 463–478. <https://doi.org/10.1093/scipol/scy074>

Vienni Baptista, B., Fletcher, I., Maryl, M., Wciślik, P., Buchner, A., Lyall, C., Spaapen, J., & Pohl, C. E. (2020). *SHAPE-ID: Shaping Interdisciplinary Practices in Europe: Deliverable 2.3: Final Report on Understandings of Interdisciplinary and Transdisciplinary Research and Factors of Success and Failure: Vol. D 2.3* [Report]. ETH Zurich.
<https://doi.org/10.5281/ZENODO.3824839>