Formative evaluation of transdisciplinary research for systematic impact orientation in real-world laboratories

There are few approaches to evaluating the societal impact of transdisciplinary research that assess and promote impact orientation while the project is still running. In this Design Report, *we present a framework for designing and conducting accompanying formative evaluation of impact-oriented transdisciplinary research in real-world laboratories. Examples from two research projects from the fields of biodiversity management and consumer logistics in rural regions illustrate the application of our evaluation design.*

Jasmin Wiefek, Emilia Nagy, Martina Schäfer

Formative evaluation of transdisciplinary research for systematic impact orientation in real-world laboratories
GAIA 33/S1 (2024): 94–101

Abstract
Scientific literature offers a variety of endeavors to conceptualize and assess societal effects of research. However, it lacks approaches on how to foster impact orientation and evaluate the effectiveness of transdisciplinary research and real-world laboratories. In this report, we present a framework for designing and conducting project-accompanying formative evaluation for impact-oriented monitoring of research processes, and for systematically recording their effects. We illustrate the application of our evaluation design with examples from two different real-world laboratories that we have been monitoring as evaluators for two years. We discuss the challenges in categorizing outputs and effects and in clarifying their significance for project success. We also address the limitations of our approach with respect to data quality and feasibility of data collection. Formative evaluation is a challenging task for all project participants. It requires adequate resources that must be considered during project planning.

Keywords
evaluation design, formative evaluation, impact-relevant indicators, real-world laboratories, societal effects, systematic impact orientation, transdisciplinary research

Dr. Jasmin Wiefek (corresponding author) | Technische Universität Berlin | Center Technology and Society (ZTG) | Berlin | DE | wiefek@ztg.tu-berlin.de

Emilia Nagy, MA | Technische Universität Berlin | Center Technology and Society (ZTG) | Berlin | DE | nagy@ztg.tu-berlin.de

Prof. Dr. Dr. Martina Schäfer | Technische Universität Berlin | Center Technology and Society (ZTG) | Berlin | DE | schaefer@ztg.tu-berlin.de

https://doi.org/10.14512/gaia.33.S1.14
Received May 12, 2023; revised version accepted December 21, 2023 (double-blind peer review).

In the realm of sustainability research, there has been notable momentum towards transformative research within real-world laboratories (RwLs). An RwL refers to a transdisciplinary (TD) research facility designed to conduct experiments in a spatially defined societal context (Parodi et al. 2016). RwLs comprise aspects of lab design, real-world or transformative experiments, and interventions (Kampfmann et al. 2023). There is a relationship among these layers, where experiments, conducted within RwLs as research environments, utilize interventions to generate evidence on sustainability solutions (Kampfmann et al. 2023) RwLs aim to initiate transformation processes and to sustain corresponding scientific as well as societal learning processes (Parodi et al. 2016). Reflection on the assumed impact pathways is pivotal for the consolidation and transferability of the results gained in RwLs (Schneidewind and Rehm 2019). We respond to the call of the scientific community for formative evaluation to ensure high-quality TD processes in RwLs (Defila and Di Giulio 2018, Bergmann et al. 2021) and methods which allow evaluation of RwLs (Schäpke and Beecroft 2022, Williams and Robinson 2020).[1] In particular, accompanying formative evaluation is supposed to enable reflexivity and learning in RwLs for impact-oriented adaptation of the research process (Bergmann et al. 2005). In this report, we present a framework for the design of project-accompanying formative evaluation and give examples of how we apply the evaluation approach in two different RwLs. We discuss our experiences to make recommendations for future use.

Sample projects

We draw on experiences from two TD research projects we accompany as evaluators (table 1). The *Biodiversity Valuing & Valuation (BioVal)* project brings together science and business in

1 For an evaluation approach that focuses on observing impact in real time, see the social design lab, which allows for process iteration and captures intangible changes (Franck et al. 2024, in this issue).

TABLE 1: Characteristics of the two sample projects *Biodiversity Valuing & Valuation (BioVal)* and *UCKER Warentakt (UWT)*.

PROJECT CHARACTERISTICS	BIOVAL	UWT
field of action (scope)	biodiversity impact assessment methods and corporate sustainability management methods	social logistics in rural regions
objectives	development and test of entrepreneurial management tools that enable companies to assess the impact of their food products on biodiversity in order to reduce the negative impacts	development of two tested logistic services for strengthening regional public transport, for improving supply in rural regions in northeastern Germany and for reducing individual automobile traffic due to the decrease in shopping trips and parcel delivery in the region
methods/participants	Scientists from the field of biodiversity impact assessment and sustainability management research are working with three companies from the food industry on developing and testing management tools that can also be transferred to other companies.	In collaboration with members from the executive board of the public transport agency, scientists from the local university of applied sciences set up an RwL along a bus route connecting a town with three villages in a rural region in Brandenburg with 14 local retailers, one parcel shop, and one small shop in each village.
project structure	consortium led by one of the academic partners; project with five modules, three of them designed as RwLs	The RwL is structurally embedded in a transdisciplinary alliance for regional innovation as a project with its own structure.
communication and knowledge integration structures	monthly steering group meeting of the scientific partners; regular meetings of the scientific and non-scientific partners; annual (reflection) workshops about: 1. progress of the project, 2. integration of knowledge, 3. intended and achieved societal impact of the project; knowledge transfer forums every six months for companies external to the project; publication of scientific papers as well as methodological and management guidelines	monthly project meetings of the scientists for operative decisions; annual meetings with the project partners for strategic questions; accompanying feedback talks with the practitioners in the testing phase on an almost daily basis for quick solutions to problems and challenges; bi-monthly meetings of the evaluator and the scientific project assistant for monitoring; three project workshops applying Theory of Change
duration of the project	3 years (11/2021 to 10/2024)	2.5 years (8/2020 to 1/2023)
embeddedness of the evaluation	One research associate who is part of the project serves as project evaluator.	Evaluator is internal to the alliance, but external to the RwL.

order to develop and implement tools for effective biodiversity management. Three scientific institutions and three companies from the food industry, all based in Germany, are involved in the project. The dissemination of results is ensured by offering knowledge transfer forums for practitioners from the food industry. In addition, project insights will be published, and a final conference is planned.

The *UCKER Warentakt (UWT)* project aims to introduce new logistics services in order to strengthen local public transport, reduce car traffic, and improve the supply of goods in a rural area in the northeast of Germany: One service delivers goods bought in stores in town, using small shops in the villages as pick-up stations. The other service delivers parcel returns from inhabitants in these villages to parcel shops in the nearby town. Both services rely on regular bus services run by the regional public transport company. The RwL invites interested retailers, parcel services, and citizens to try out the new services. The process is supported by scientists.

Both projects are classified as RwLs due to their transdisciplinary and experimental nature. Their primary goals are to contribute to sustainability transformation and experimenting with potential solutions. Notably, the communication and knowledge integration structures as well as the learning environment are designed for long-term perspectives and the transferability of results. The inclusion of formative evaluation enhances their reflexivity, aligning with the inherent characteristics of RwLs (Schäpke et al. 2018).

Methodological approach

Our formative evaluation runs parallel to the RwLs and is based on Theory of Change (ToC) as analytical framework (Schneider et al. 2019). The approach involves to formulate a shared vision in a group of different stakeholders and to plan, monitor, and assess processes, outputs, and effects (Belcher et al. 2020). When we conduct evaluation, we pursue the following goals: One objective is to assess the success of the project. Referring to the multilayer evaluation approach of RwLs by Kampfmann et al. (2023), at the level of interventions, the aim is to determine immediate outputs and effects. At the level of experiments, the focus is on internal aspects like the quality of the collaboration and >

impact-oriented adaptation of processes. And at the level of lab context, the evaluation emphasizes improvements in the future and sustaining the results. The evaluation can take on a strong formative character when the evaluator primarily assists the project team in evaluating whether they are achieving their project goals. The evaluation may assume a somewhat different character when evaluators may (in part) conduct assessments themselves with a focus on performance measurement (Kuhlmann 2003).

Further, we aim to improve the impact-oriented performance of the projects. We regularly reflect with the project partners on assumptions about how societal impacts can be achieved through activities and outputs initiated by both scientific and non-scientific actors. These collaborative reflection processes within the project provide insights into necessary adjustments to the project activities. The evaluation, therefore, assumes a formative character, focusing on the ongoing project as a learning objective while also intending to impart lessons for (the design of) impactful future projects (Bergmann et al. 2005). Moreover, the observations obtained from the evaluation aim to provide insights into generalized impact pathways of TD research. Thus, in addition to the scientific outputs primarily generated in the project under evaluation, the evaluation itself serves as a scientific output.

Designing formative evaluation

Theory of Change

The initial step in our evaluation design is to develop a project-specific ToC. At the beginning of the project, we conduct an "impact workshop" to jointly describe impact pathways towards a shared vision[2]. The project participants reflect on the project structure, objectives, and planned integration steps and negotiate the following questions: *What is the shared project vision? What are the main project activities and interactions between different actors? Who are the key actors in the project context and what are their roles? Which outputs does the project aim for? Which societal effects are intended and to what extent are they expected to occur during and beyond project duration?* Intended and possible side effects are sorted and visualized using the heuristic for systematizing societal effects of TD research (Schäfer et al. 2021). According to the heuristic, effects are classified into three orders: 1st order effects are expected to occur within the duration or spatial scope of the project; 2nd order effects are expected to occur after the project is finished, but within the close temporal or spatial context of the project; 3rd order effects are expected to occur beyond the temporal and spatial context of the project. Connecting the collected elements of activities, outputs, and effects creates a ToC with pathways that may also involve positive or negative feedback loops (figure 1). In addition, participants discuss the extent to which they can control or merely influence the occurrence of effects (Belcher et al. 2020).

Monitoring concept

The collected components of the ToC are the basis for the second step in conducting the evaluation: the development of a monitoring concept. Table 2 (pp. 98/99) shows a monitoring concept with examples from *BioVal* and *UWT*.

For the monitoring concept, we define project-specific evaluation *criteria* together with the project team. Criteria refer to abstract concepts against which to reflect and evaluate the project. Besides criteria for the outputs and effects, we also include criteria for the quality of the research process, based on the assumption that the quality of the research process has an influence on the effectiveness of the project. Impact-relevant process qualities of TD research refer, for example, to the collaboration and communication culture within the project (Lux et al. 2019, Williams and Robinson 2020).

Once the criteria for the quality of the research process, outputs, and effects have been formulated, project-specific, quantitative and qualitative *indicators* have to be defined for each criterion. In formative evaluation, the project team collaboratively agrees on a set of indicators. When developing indicators, the following points should be considered: 1. the (empirical) data that "feeds" an indicator must be available or collectable; 2. it must be possible to collect the data during the project period; 3. the project participants must have the necessary capacity to collect the data; and 4. the project participants take responsibility for data collection.

In the next step, the operationalization, the project team specifies the required data. The difference between indicators and operationalizations is that indicators describe the necessary information that must be collected in general terms. This can be, for example, that the project participants *think* that their perspectives are considered in an appropriate manner (e.g., in decision-making). Operationalization is the process of transforming this information into measurable constructs (Mörtel et al. 2023). In the example given, this would be that the project participants *state* that their perspectives were appropriately considered in decision-making. This step allows to determine the specific *methods for data collection,* such as interviews, questionnaires, participatory observation or documentation of workshops. For the example above, we have chosen a survey to ask the project participants about the aspects of interest ("Please tick on a scale of 1 to 10: My views have been adequately considered in decision-making").

Following this procedure, a monitoring concept is developed which also documents who is *responsible for collecting the specific data*. It is important to clarify who is responsible for a) the conceptual preparation of the monitoring materials (questionnaires, documentation templates, etc.); b) the data collection; c) the data analysis; d) facilitating the joint reflection of the evaluation results with all project members. It is necessary to distribute the responsibilities for the different tasks (a to c) between the proj-

2 The method of conducting a workshop, involving impact reflection and the development of plausible impact chains using a Theory of Change approach, was initially developed in the tdAcademy project by Josefa Kny and Martina Schäfer. In the *BioVal* and *UWT* projects, this workshop method was integrated as one component of the broader monitoring concept.

ect participants. Usually, the evaluator is responsible for the task of facilitating the joint reflection of the evaluation results (d).

Data collection, joint reflection, and impact-oriented readjustment

During the course of the project, *data is continuously collected and analyzed* with the help of the monitoring concept. One important step in the accompanying formative evaluation is that the results are regularly discussed with the project team. This enables joint decisions to be made on whether the project's processes are working well and whether the project is on track to achieve its intended outcomes and effects. Repeated *reflection on the ToC* provides a good setting for discussing *strategic readjustments* to the project design. This reflection process also helps to identify risks to sustaining implemented solutions and intended effects beyond the duration of the project (Mayne 2020). This *risk analysis* is particularly important at the end of the funding period, before the scientific partners leave the project team.

Figure 2 (p. 100) summarizes the elements we consider relevant for designing and conducting a formative evaluation.

Discussion

In the following, we discuss some of the issues that we consider to be particularly relevant to a successful formative evaluation or where we see potential for improvement of future application.

Significance and categorization of outputs and effects

In *BioVal* and *UWT,* the project participants could follow the theoretical considerations of process evaluation (e.g., that the quality of the co-production and communication processes influence a project's success and effectiveness). Nevertheless, in *UWT,* for example, some practitioners focused strongly on outputs, such as the number of parcels delivered. Our evaluator's perspective that the joint set-up of the new services generates important societal effects such as learning processes and network-building initially received little attention. Neglecting the relevance of these effects threatened the long-term establishment of the new services. Important processes such as building competencies for a seamless service provision or the search for further partners were not well-nurtured from the beginning. A common understand- >

FIGURE 1: Sample Theory of Change (ToC) (based on Belcher et al. 2020, Schäfer et al. 2021). Based on the project participants' discussions, activities, outputs and effects are linked and a ToC with impact pathways is created. These pathways lead to a shared vision and may have positive or negative feedback loops. The decreasing line thickness is intended to symbolize the diminishing influence on the effects.

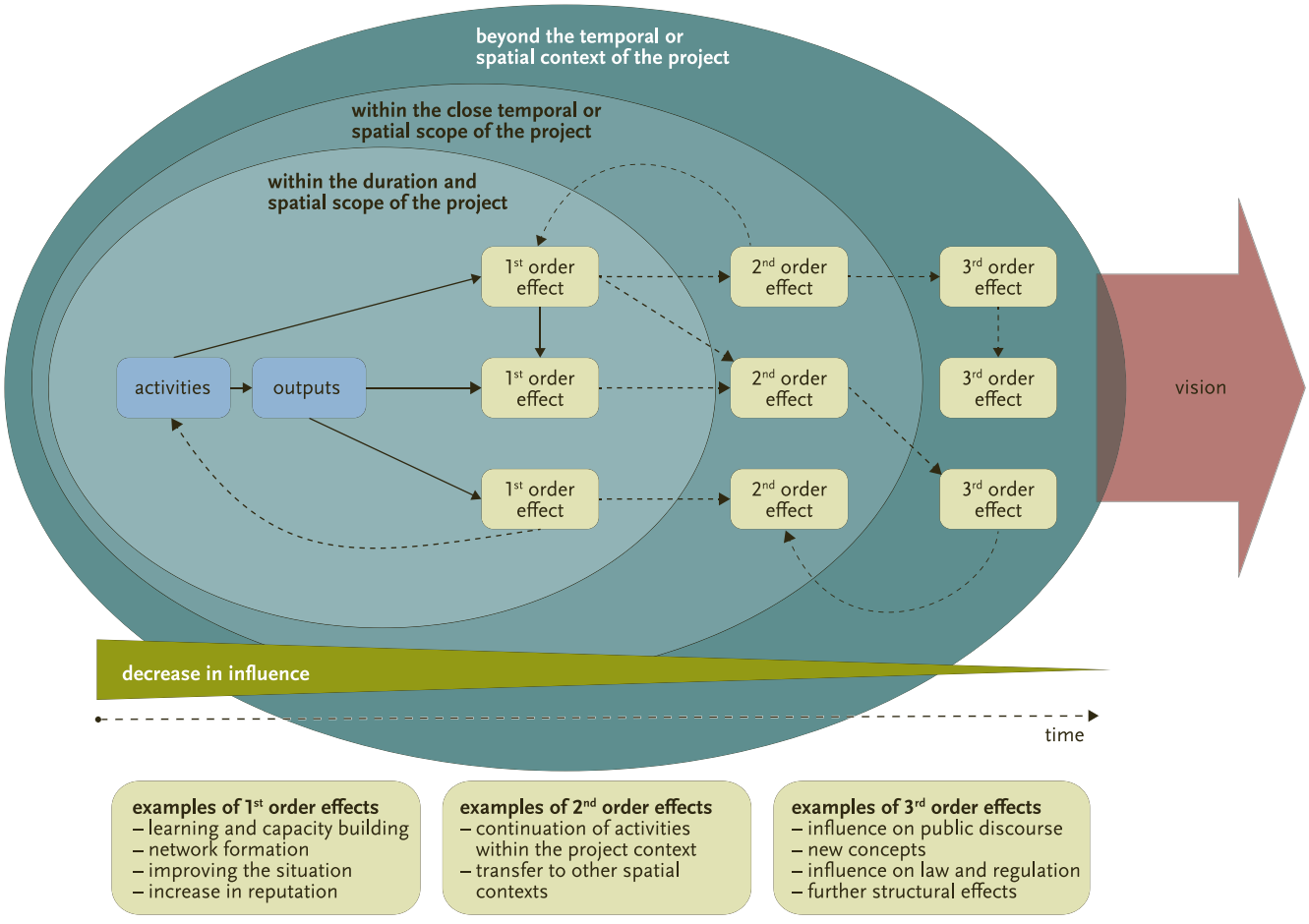

TABLE 2: Sample monitoring concept with examples from *Biodiversity Valuing & Valuation (BioVal)* and *UCKER Warentakt (UWT)*. a: preparation of the monitoring materials, b: data collection, c: data analysis, d: facilitating the joint reflection of evaluation results.

CRITERIA	INDICATORS
PROCESS QUALITIES	
participation	The project participants think that their perspectives are included in an appropriate manner.
	Sufficient members of all relevant stakeholder groups participate actively in the co-production process.
reflection and iterative adjustment	The partners develop a shared understanding of the intended effects of the project, which they repeatedly reflect on and, if necessary, adapt during the course of the project.
	The project participants can identify with the ToC and consider it a good basis for the readjustment of the research design.
OUTPUTS	
Biodiversity protection is part of the sustainability management of the RwL companies. *(BioVal)*	Biodiversity protection is integrated into the sustainability management of the three RwL companies.
New logistics services with their organizational elements are successfully tested. *(UWT)*	On a test route, parcels were delivered from the city to customers in the villages and vice versa by public buses.
1ST ORDER EFFECTS	
learning and capacity building *(BioVal)*	The awareness for the protection of biodiversity is increasing among the employees in the RwL companies during the *BioVal* project runtime.
improving the situation *(UWT)*	In the test region, individual traffic by car is decreasing.
2ND ORDER EFFECTS	
transfer to other spatial contexts *(BioVal)*	Companies from the transfer forum intend to integrate the protection of biodiversity into their sustainability management.
continuation of activities within the project context *(UWT)*	The new services will continue under the new business model of the transport company and will be extended to more bus routes.
3RD ORDER EFFECTS	
influence on laws and regulation *(BioVal)*	Introduction of a legal requirement that biodiversity protection must be integrated into the sustainability management of companies.
influence on public discourse (UWT)	Public transport companies are perceived as important actors promoting socio-ecological transformation in the field of parcel logistics.

ing of the significance of these effects was achieved by analyzing the risks to the continuation of the services. This analysis led to corresponding actions. This example from *UWT* shows that an early discussion about the relevance of outputs and effects on the project's success would have been fruitful. We also recommend paying attention to what the project partners define as project goals, outputs, and effects. The question of whether network formation is understood as an output or an effect, for example, is a project-specific, case-by-case decision that may need to be discussed repeatedly over the course of the project. The categorization of the examples of societal effects in figure 1 should thus be understood as variable rather than fixed.

Embeddedness of the evaluation and associated roles

In conducting the formative evaluation, the evaluators and other project participants take on specific roles. The tasks of the

OPERATIONALIZATION	METHODS FOR DATA COLLECTION	RESPONSIBILITIES
The *BioVal* partners indicate that their perspectives were appropriately considered in the steering committee.	post-event feedback questionnaire	a, b, c, and d: evaluator
At least one representative from each stakeholder group, considered relevant for the process by the project team, participates at one event of co-production or more.	analysis of lists of participants of co-production events	a, b, c, and d: evaluator
At the beginning, in the middle and towards the end of the project, an "impact workshop" is held in which the project participants clarify what effects they expect from the project and which readjustments are necessary in the process.	documentation of impact workshops	a, b, c, and d: evaluator
When giving feedback on the impact workshops, the participants state that they can identify with the ToC and consider it a good basis for the adjustment of the research process.	documentation of feedback rounds in the impact workhops	a, b, c, and d: evaluator
At the end of the project, the representatives of the three RwL companies state that biodiversity protection is integrated into their sustainability management.	interviews and statements from the final impact workshop	a, b, c, and d: evaluator
A minimum of parcels (10/month) was delivered along the test route in both directions.	documentation of delivered parcels	a, b, and c: project coordinator and scientific assistant RwL; d: evaluator
The employees in the RwL companies indicate how highly they rate their awareness of the need to protect biodiversity at two points in time (project start and end).	online survey	a c, and d: evaluator; b: project members of the RwL companies
A significant number of customers (at least 10) claim that they were able to forgo a trip into town because the parcel was delivered to them.	interviews	a, b, and c: project coordinator and scientific assistant RwL; d: evaluator
Representatives from companies attending the transfer forum state that their companies intend to integrate the protection of biodiversity into their sustainability management.	online survey and interviews	a, b, c, and d: evaluator
The stakeholders involved in the implementation of the services sign a cooperation agreement in order to maintain the services after the end of the project period.	documentation if the cooperation agreement was signed	a, b, and c: project coordinator; d: evaluator
Attribution of effects to single projects is no longer possible if effects occur in spatial and temporal distance. Since clear impact pathways between project activities and 3rd order effects can rarely be drawn, it is only possible to collect plausible hints for these links.		

evaluators can be described with several roles of an integration expert according to Hoffmann et al. (2022). In both our sample projects, we, as *evaluators,* had the role of *translators,* bringing together the different perspectives of the project participants and asking questions about the impact mechanisms for discussion. We were also *facilitators,* since we facilitated knowledge integration processes about the evaluation findings. At the same time, we have been researching approaches to evaluating and designing impactful research. Therefore, we are also *contributors* to this field of research. Conducting a formative evaluation requires not only the expertise of the evaluators in these different roles, but also the capacities of the other project participants. They are both data providers, for example, by filling in questionnaires, and active players in data collection, for example, by documenting workshops. Moreover, they participate in evaluation workshops, reflecting on the project process and evaluation results. It must be >

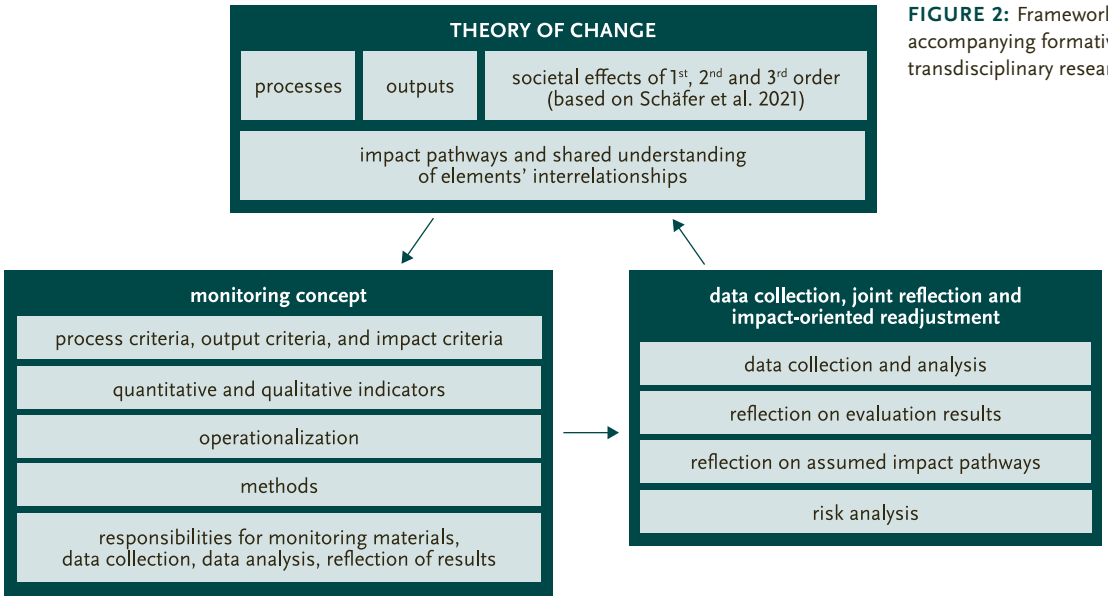

FIGURE 2: Framework for the design of an accompanying formative evaluation of transdisciplinary research.

clearly discussed with the project participants and considered in the project design that formative evaluation involves the efforts of all project participants.

Data collection

The definition of indicators implies that the expected outcomes and effects are empirically observable. However, since the practicability of data collection plays a major role in the development of indicators, this claim cannot always be realized, especially with regard to higher-order effects. In *BioVal,* for example, the aim is to contribute to biodiversity conservation in the long term by providing a biodiversity management tool. Whether this effect will occur cannot be predicted or documented in the course of the project. Vice versa, improved protection of biodiversity in the future cannot be attributed to the activities of the *BioVal* project, as the project is one actor among others in this field. The extent to which, or the conditions under which, the occurrence of these effects becomes more likely can only be estimated with the help of a plausibility and risk analysis of the impact pathways. Similarly, for capturing effects such as an increase in awareness of biodiversity conservation among employees in the RwL companies, collecting data at least at two points in time would be optimal: a baseline survey at the beginning of the project – or even before the project starts – and a second survey at the end. However, due to short project durations (two to three years) these two surveys might be scheduled too close to each other to capture any changes. In order to capture such long-term effects, it would be necessary to carry out surveys a few years after the end of the project. However, it has to be considered whether the expected range of effects justifies the expense of long-term impact monitoring.

Data quality

One limitation of the evaluation is the quality of the available data. For example, the RwL companies from the *BioVal* project are expected to consider biodiversity concerns in their decisions after the successful integration of biodiversity issues in their sustainability management. Whether or not this is the case will be ascertained towards the end of the project via interviews with one representative from each company. In this respect, the evaluation is based on self-reported data or statements by individuals. A validation of these statements would generally be conceivable, for example, by interviewing further company employees or analyzing company documents. In *UWT,* for example, we used questionnaires and participatory observation as complementary methods to assess process criteria such as transparency. Yet, using more than one method implies higher costs. There is therefore a certain tension between the realization of a manageable evaluation process and the validity of the data. Nevertheless, evaluators should strive to increase the validity of the data by using mixed methods.

Conclusion

Academics, practitioners, and research funders are increasingly interested in obtaining evidence that TD research is fulfilling its claim to contribute to solving societal problems. More and more projects are integrating formative evaluation to increase the effectiveness of their research processes and to document their effects. Funding agencies provide resources for evaluation, as in our two example projects. This is a positive trend that needs to be encouraged if evaluation of TD research and RwLs is to become more important in the future, also on the part of funders. Anticipating and tracing impact pathways is a challenging endeavor which demands additional personnel and time resources. For larger projects or programs, it might be particularly informative to combine formative evaluation, conducted concurrently, with summative or ex-post evaluation taking place one, two, or five years after the respective research process.

In the scientific discourse, there are different endeavors to conceptualize societal effects (Augenstein et al. 2022, Marg et al. 2019, Moser and Wolf 2023). In our view, discussion around concepts, approaches, criteria, and methods is particularly important. Criteria and indicators for TD projects are context and case-specific; nevertheless, it would be worth examining whether a standardization of the evaluation of TD projects, for example, through a list of generally applicable process criteria and manuals for planning the evaluation, would facilitate the work of evaluators. As a first step in this direction, our report shares experiences of designing and conducting project-accompanying formative evaluation of TD RwLs.

Acknowledgement: We would like to thank *Josefa Kny* and two anonymous reviewers for their valuable comments on this paper.
Funding: This work was supported by the German Federal Ministry of Education and Research (BMBF). Funding codes: 03WIR0805A, 03WIR0805, and FKZ 01UT2110A–F.
Competing interests: The author(s) declare no competing interests.
Author contribution: *JW, EN:* framework and monitoring concept development, collection and analysis of data, manuscript drafting and writing the final manuscript, creating visualizations; *EN:* contribution to funding acquisition. *MS:* funding acquisition, supervision, support in framework and monitoring concept development, support in analysis of data, reviewing and editing manuscript draft and final manuscript.

References

Augenstein, K., P. M. Bögel, M. Levin-Keitel, H. Trenks. 2022. Wie entfalten Reallabore Wirkung für die Transformation? Eine *embedded-agency perspective* zur Analyse von Wirkmechanismen in Reallaboren. *GAIA* 31/4: 207–214. https://doi.org/10.14512/gaia.31.4.4.

Belcher, B. M., R. Davel, R. Claus. 2020. A refined method for theory-based evaluation of the societal impacts of research. *MethodsX* 7: 100788. https://doi.org/10.1016/j.mex.2020.100788.

Bergmann, M. et al. 2005. *Qualitätskriterien transdisziplinärer Forschung. Ein Leitfaden für die formative Evaluation von Forschungsprojekten.* ISOE-Studientexte 13. Frankfurt am Main: Institut für sozial-ökologische Forschung (ISOE). www.isoe-publikationen.de/fileadmin/redaktion/ISOE-Reihen/st/st-13-isoe-2005.pdf (accessed January 3, 2024).

Bergmann, M. et al. 2021. Transdisciplinary sustainability research in real-world labs: Success factors and methods for change. *Sustainability Science* 16/2: 541–564. https://doi.org/10.1007/s11625-020-00886-8.

Defila, R., A. Di Giulio (Eds.). 2018. *Transdisziplinär und transformativ forschen: Eine Methodensammlung.* Wiesbaden: Springer Nature. https://doi.org/10.1007/s11625-020-00886-8.

Franck, M., N. Hempel, S. Preiß, R. Boch. 2024. Creating and observing impacts in transdisciplinary projects – Insights from the social design lab. *GAIA* 33/S1: 110–115. https://doi.org/10.14512/gaia.33.S1.16.

Hoffmann, S., L. Deutsch, J. T. Klein, M. O'Rourke. 2022. Integrate the integrators! A call for establishing academic careers for integration experts. *Humanities and Social Sciences Communications* 9/1: 147. https://doi.org/10.1057/s41599-022-01138-z.

Kampfmann, T., P. Bernert, D. J. Lang. 2023. Toward a modular evaluation approach of real-world laboratories: Findings from a literature review. *Research Evaluation* 32/1: 128–143. https://doi.org/10.1093/reseval/rvac029.

Kuhlmann, S. 2003. Leistungsmessung oder Lernmedium? Evaluation in der Forschungs- und Innovationspolitik. *TATuP – Zeitschrift für Technikfolgenabschätzung in Theorie und Praxis* 12/1: 11–19. https://doi.org/10.14512/tatup.12.1.11.

Lux, A. et al. 2019. Societal effects of transdisciplinary sustainability research: How can they be strengthened during the research process? *Environmental Science and Policy* 101: 183–191. https://doi.org/10.1016/j.envsci.2019.08.012.

Marg, O., M. Kreß-Ludwig, A. Lux. 2019. *Wirkungen transdisziplinärer Stadtforschung in den Projekten der Förderlinien „Leitinitiative Zukunftsstadt" und „Nachhaltige Transformation urbaner Räume". Wirkungskategorien, Projektprofile und Handreichung zur Selbstreflexion. Werkstattbericht Synthese- und Vernetzungsprojekt Zukunftsstadt SynVer*Z.* www.nachhaltige-zukunftsstadt.de/downloads/SynVerZ_2019_Werkstattbericht_Wirkungen_transdisziplinarer_Stadtforschung.pdf (accessed January 3, 2024).

Mayne, J. 2020. Sustainability analysis of intervention benefits: A Theory of Change approach. *Canadian Journal of Program Evaluation* 35/2: 204–221. https://doi.org/10.3138/cjpe.70004.

Mörtel, J., A. Nordmann, O. Schlaudt (Eds.). 2023. *Indikatoren in Entscheidungsprozessen. Stärken und strukturelle Schwächen.* Wiesbaden: Springer VS. https://doi.org/10.1007/978-3-658-40638-7.

Moser, A., B. Wolf. 2023. Wirkungspotenziale von Forschung und Entwicklung bewerten: Ein theoriebasiertes und multidimensionales Kriterien-Set. In: *Indikatoren in Entscheidungsprozessen. Stärken und strukturelle Schwächen.* Edited by J. Mörtel, A. Nordmann, O. Schlaudt. Wiesbaden: Springer VS. 57–68. https://doi.org/10.1007/978-3-658-40638-7_10.

Parodi, O. et al. 2016. Von „Aktionsforschung" bis „Zielkonflikte". *TATuP – Zeitschrift für Technikfolgenabschätzung in Theorie und Praxis* 25/3: 9–18. https://doi.org/10.14512/tatup.25.3.9.

Schäfer, M., M. Bergmann, L. Theiler. 2021. Systematizing societal effects of transdisciplinary research. *Research Evaluation* 30/4: 484–499. https://doi.org/10.1093/reseval/rvab019.

Schäpke, N., R. Beercroft. 2022. From control to co-production: Eight steps to monitor, evaluate, and adapt participatory experiments. *Social Innovations Journal* 15/5. https://socialinnovationsjournal.com/index.php/sij/article/view/5038 (accessed January 3, 2024).

Schäpke, N. et al. 2018. Jointly experimenting for transformation? Shaping real-world laboratories by comparing them. *GAIA* 27/S1: 85–96. https://doi.org/10.14512/gaia.27.S1.16.

Schneider, F. et al. 2019. Transdisciplinary co-production of knowledge and sustainability transformations: Three generic mechanisms of impact generation. *Environmental Science and Policy* 102: 26–35. https://doi.org/10.1016/j.envsci.2019.08.017.

Schneidewind, U., A. Rehm. 2019. Vom *inside-out* zum *outside-in:* Perspektivwechsel bei der Impact-Messung von transformativer Forschung. *GAIA* 28/2: 168–170. https://doi.org/10.14512/gaia.28.2.18.

Williams, S., J. Robinson. 2020. Measuring sustainability: An evaluation framework for sustainability transition experiments. *Environmental Science and Policy* 103: 58–66. https://doi.org/10.1016/j.envsci.2019.10.012.

Jasmin Wiefek
Diploma (University of Göttingen, DE) and PhD (Freie Universität Berlin, DE) in psychology. Since 2012, research associate in socio-ecological transformation research. Research interests: sustainable consumption, sustainable business, evaluation of transdisciplinary research.

Emilia Nagy
Diploma in physics (Eötvös University Budapest, HU) and MA in German literature, linguistics, art, and media sciences (Universität Konstanz, DE). Since 2012, research associate in transdisciplinarity. Research interests: methods for inter- and transdisciplinary research, societal effects of transdisciplinary research.

Martina Schäfer
PhDs in environmental technology and sociology (Technische Universität Berlin, DE). Since 2010, scientific director of the Center for Technology and Society, Technische Universität Berlin. Research interests: sustainable consumption, sustainable regional development, methods for inter- and transdisciplinary research.