SOCIAL INNOVATIONS Journal Vol. 15 (2022): Sustainability Transformations in Practice

From Control to Co-production: Eight Steps to Monitor, Evaluate, and Adapt Participatory Experiments

By: Niko Schäpke*, Richard Beecroft*

*Assistant Professor, Faculty of Environment and Natural Resources, University of Freiburg, Germany
*Managing Director, Center Humans and Technology, Karlsruhe Institute of Technology

Corresponding author: Niko Schäpke, niko.schaepke@ifp.uni-freiburg.de

Keywords: Transdisciplinarity, Real-world labs, Evaluation, Interventions, Sustainability Transformation

Abstract

Co-creative and action-oriented sustainability research, including real-world labs, (urban) living labs, and transformation labs, arose from the desire to contribute to societal transformations. Mentioned labs use experiments to test ideas for a more sustainable life and to promote changes toward sustainability on the ground. As social and scientific actors implement participatory experiments together, social engagement is central to their success. It is important to keep an eye on the impact of an experiment while going along. Monitoring and evaluation allow adjustments at an early stage of experiments. In addition, data can be collected for final evaluation of whether a participatory experiment was successful and why and if it can be transferred or duplicated. Overall, an important societal and scientific learning opportunity is created. While highly promising, such monitoring and evaluation is a challenging task. It depends on continuous interaction with stakeholders for data collection and reflection. Attentive monitoring and adaption might strengthen stakeholder engagement and vice versa. They could even be integrated into the overall co-creation of participatory experiments and labs. Yet, this fruitful interaction has to be worked for, requiring delicate decisions and practical know-how.

This contribution is oriented towards supporting practical applications. It outlines eight steps of how to design, plan and implement the monitoring of a participatory experiment: 1) agree on the objective; 2) determine the experiment and monitoring scope; 3) determine the parameters and indicators of measurement; 4) determine the timing, type and medium of data collection; 5) collect and store data; 6) analyze and evaluate the data; 7) present and communicate the results; 8) adapt the experiment. Steps have a cyclical, iterative nature. Both an ideal-type monitoring scheme and a plan are presented to guide application. A productive interrelation of monitoring and facilitating engagement is discussed and illustrated based on a practical example.

Introduction

Co-creative and action-oriented sustainability research arose from the desire to contribute to necessary socio-ecological-technical transformations (Fazey et al., 2018; Caniglia et al., 2020; Norström et al., 2020). This includes transdisciplinary and experimental approaches such as real-world labs, (urban) living labs, and transformation labs (Fazey et al., 2018;

McCrory et al., 2020). These approaches, called labs hereafter, use experiments to test ideas for a more sustainable life and to promote changes toward sustainability on the ground (Schäpke et al., 2018; Caniglia et al., 2017). As social and scientific actors often implement interventions together, social engagement is central to the success of such participatory experiments[i].

It is important to keep an eye on the impact of a participatory experiment while it is being implemented (Luederitz et al., 2017; Williams and Robinson 2020; Beecroft et al., 2018). Through monitoring and evaluation, experiments can be adapted at an early stage, and the effects on society can be readjusted. Simultaneously, data can be collected for a final evaluation, creating an important learning opportunity (Luederitz et al., 2017). Data allows conclusions about the extent to which an intervention was successful, why, and if it can be transferred or duplicated (Lam et al., 2020; Beecroft et al., 2018). While highly promising, setting up and implementing such monitoring and evaluation is not an easy task. It draws on information and perspectives gained in continuous exchange and reflection with stakeholders (Beecroft et al., 2018). Although attentive monitoring and adaption might strengthen engagement and support, this fruitful interaction has to be worked for. It appears supportive of data collection, and assessment is done to benefit stakeholders' interests and capacities. Thus, the appraisal and adaptation of a participatory experiment, as well as its co-creation, are interrelated and might fruitfully be combined.

The present contribution is oriented towards guiding practical application. It draws on the authors' experiences with various participatory experiments, labs, and scholarly publications. The contribution has the following structure: First, it outlines eight steps of how to design, plan and implement the monitoring, evaluation, and adaption of a participatory experiment. This includes the presentation of an ideal-type monitoring and evaluation scheme and a monitoring and evaluation plan. Second, it discusses a productive interrelation between monitoring and facilitating engagement. Third, it presents an application example based on an international university living-lab collaboration, outlining tools at the intersection of monitoring and evaluation and facilitating participation with varying intensities.

Design, planning, and implementation of monitoring

The process of continuous observation[ii] is outlined below. The steps presented are ideal-typical and must be adapted to the respective circumstances. In principle, steps should be concretized and implemented by a team of practitioners and scientists, allowing monitoring measures to be broadly supported and insights to be shared. In case of overly scarce resources, concretization and implementation could also be done by a single person, such as lab management. Monitoring has a political dimension when it concerns assessing the success of an experiment. This has implications for the experiment's continuation and stakeholder activities. Conflicts of interest and tensions can arise. To avoid these, the monitoring aims and procedures should be transparent and, ideally, be co-created and agreed upon with the participants.

Steps 1-4, the conceptualization, and planning of the monitoring, should generally be taken before the intervention begins. *Steps 5-8* concern the implementation of the monitoring and the evaluation as well as the adaptation of the experiment.

1. **Agree on the objective of the monitoring:** Is it to obtain data and scientifically analyze how the experiment works? Or is it to ensure the support of the participants? Or both? Should monitoring be used to promote participation? In addition to the objective in terms of content, the available resources (time, personnel, and finances) should also be considered from the outset. Sound, shared, and transparent objectives are crucial for the acceptance and therewith potential success of monitoring and evaluation efforts. Thus, this step requires caution and sufficient time investment and may include intense deliberations and even temporary conflict.

2. **Determine the experiment and monitoring scope:** What exactly is considered an experiment? This can be, for example, technical, social, or legal changes or combinations thereof. In relation to which areas should the process and effect of the experiment be assessed? Where are the limits of the observation, e.g., in time and space? Here, areas are determined which are to be covered and considered. These can be thematic areas or phases of the experiment. A basic distinction includes four dimensions:
    a. Inputs: What is invested in the experiment?
    b. Processes: Which processes are implemented with it?
    c. Outputs: What is produced/implemented directly by the experiment?
    d. Effects: What other effects can be observed?

    For the selected areas, more specific aspects can then be described to be considered. For example, which aspects are important inputs for the experiment, its success or failure? How different actors participate in and continuously support the experiment is almost always crucial. It is important to consider not only the intended outputs and effects, but also those areas where negative effects are conceivable.

    An example of a generic monitoring scheme, including monitoring dimensions and related thematic areas, is shown in Figure 1. Here, the evaluation dimensions are understood as interconnected (Luederitz et al., 2017). The inputs to the experiment feed its processes, producing the direct results, the outputs of the experiment. Outputs, in turn, may create wider effects (see blue arrows in Figure 1). The scheme is designed to think from the aspired result of the experiment (see numbering in figure 1): I: What outputs are to be created? II: What effects are to be accomplished? Which process (III) and inputs (IV) are therefore necessary?

    The exemplary evaluation scheme indicates features for all four dimensions identified as important for the success of participatory experiments. These aspects can be used as inspiration for the development of the monitoring scheme (step 3) of any (sustainability-related) experiment: Which of the features depicted are central to your experiment? Which other aspects are relevant from your point of view? The selected features can be defined in more detail for concrete monitoring and provided with a reflection or evaluation question. For example, the aspect “trust of stakeholders/participants” can be determined by the question, “What percentage of respondents fully or at least partially supports the planned experiment?” (see Luederitz et al., 2017 for further evaluation questions).

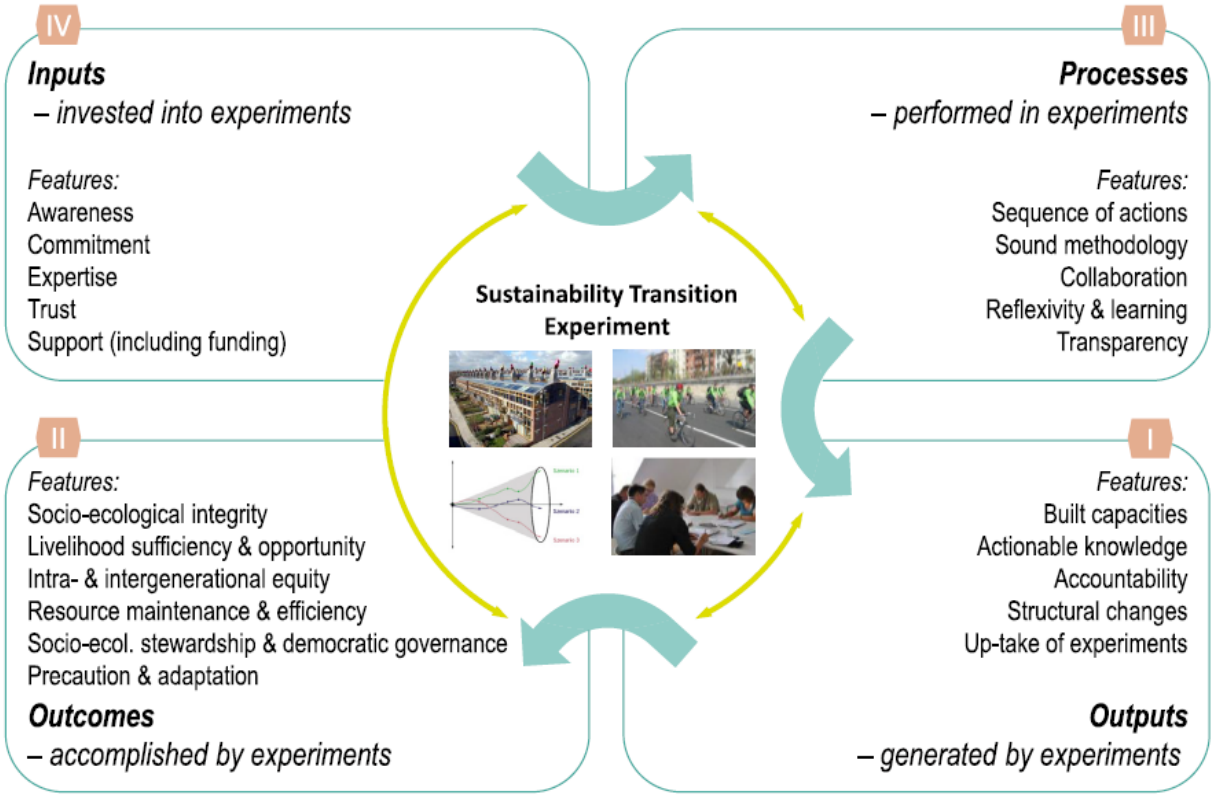

Figure 1: Generic example of a monitoring scheme, outlining dimensions and features to appraise a participatory sustainability transitions experiment. Reprinted in modified form from: Learning through evaluation – A tentative evaluative scheme for sustainability transition experiments, C. Luederitz, N. Schäpke, et al., Journal of Cleaner Production, Volume 169, 2017, Pages 61-76, Copyright (2017), with permission from Elsevier

3. **Determine concrete measurement parameters and indicators:** These make the relevant areas for the respective experiment tangible. An example in the technical field is the concrete savings in electricity or water consumption due to the experiment. An indicator is a measurable quantity that can represent an aspect, even if it does not represent it completely (e.g., indicator '% of cafes with plant milk option' for the aspect 'acceptance of vegan nutrition'). For some features, it may not be possible to name a quantitative measure or only a qualitative one. Nevertheless, try to specify how you want to evaluate the development of an aspect. The reflection and discussion about this have a learning value in themselves. With step 2, this creates a tangible monitoring and evaluation scheme consisting of monitoring dimensions, features, as well as measurements and indicators.

4. **Determine the timing, type, and medium of data collection:** Depending on the objective, available resources, and scope of observation, this can be simple or rather complex. Typical instruments for recording effects on social and participatory aspects are questionnaires, in-depth interviews, moderated group discussions (focus groups), online surveys, or even site visits. Further methods include personal meetings, discussions with

key players, and networking and communication meetings. Technical data can be represented by measurements, e.g., of energy or material balances, and economic aspects largely by common parameters such as balance data. Ecological parameters usually have to be collected specifically, e.g., by measuring pollutants or counting species to determine biodiversity. Systematic reflection on the experiments by the participants is another possible way of collecting data. Summarize in a *monitoring plan* who, how, when, and where an investigation is done (see table 1 for an example).

An important question is whether a baseline survey of the initial conditions should and could be made *before* the start of the experiment. Rigorous identification of cause and effect is often difficult in experiments under hard-to-control conditions of the 'real' world. Even without a baseline survey, ongoing, recurrent observation provides important conclusions. In this context, it is particularly important to consider the time dependence of effects, meaning the time needed for effect to become observable, which can vary significantly in different areas depending on the experiment.

Table 1: Example of a monitoring evaluation plan (own representation). Needs concretization according to the objectives of both experiment and monitoring, as well as their spatial and temporal positioning.

Who: Implementer	How: Methods	Whom: Target group				When: Quarter					Where	
		Employers	Entrepreneurs	Citizens	Technical staff	1	2	3	4	5	On site	Online
Project management	Moderated group discussions	■		■		■				■	■	
Ph.D. student	In-depth interviews		■		■		■				■	■
Web designer and project management	Online survey on project support	■	■	■	■	■		■		■		■
Technical management	Consumption data collection	■	■		■	■	■	■	■	■	■	■
Etc.	Etc.											

5. **Collect and store data to implement the monitoring plan:** This step can be carried out repeatedly, possibly also in conjunction with a timely analysis of the data (step 6) and the adaptation of the study (steps 1-4) or even the entire experiment. It is important to obtain the consent of the persons from whom data are collected (e.g., interview partners, respondents) and to document that data are collected, stored, and used. Whenever possible, data should be collected pseudonymously, i.e., names should be replaced by a small code so that answers from different points in time can be compared without making the personal reference visible. The list of codes can be destroyed at the end of the monitoring to finally anonymize the data. The further use of non-personal data may also be restricted or subject to approval (e.g., balance data of companies). The documentation of the process and experiment results is important for the evaluation. Here, it is important to consider how to ensure the feasibility of this documentation, e.g., in terms of time and resources required. Keep in mind that the results need to be evaluated, too. Consider this when deciding what data to collect and in which form (e.g., abstain from collecting data just to obtain it; use easy-to-process data formats).

6. **Analyze and evaluate the data:** Types and methods of data analysis are numerous, and these depend on the research question and the object under consideration. A simple analysis first involves the compilation and description of the data, sorted according to the various dimensions and aspects of monitoring. Then the data are discussed and evaluated individually or in a team to assess whether the intervention has been successful (so far), obstacles and challenges, and success factors. Non-intended effects should also be taken into account. More advanced forms of analysis can, for instance, include statistical analysis of quantitative data, as well as interpretative analysis of qualitative accounts.

7. **Present and communicate the results:** It is necessary to present, communicate and reflect on the monitoring results in an understandable and comprehensible way, especially if not all participants in the experiment are involved in the monitoring. This makes it possible to learn together from the results and create acceptance for possible adjustments. Careful selection of communication formats is essential if the results are critical and the support of central stakeholders is to be secured. On the one hand, communication formats should leave enough space to do justice to the results when presenting them. On the other hand, there should be room for joint discussion, possible criticism, the development of shared assessments, and the necessary adjustments to the experiments.

8. **Adapt the experiment:** Based on the knowledge gained, an adaptation of the experiment may be useful or even necessary. Reasons for an adjustment can be, for example, the decreasing support of the experiment by relevant persons/groups or the emergence of unintended side effects, which may be of a technical, economic, or ecological nature. When designing an experiment, time points can already be planned at which directional decisions are to be made based on the monitoring - this makes follow-up control considerably easier (e.g., a prototype phase followed by a selection of experiments to continue).

Subsequently, the adapted experiment can become the object of a renewed monitoring attempt. It is also possible that an entirely new experiment is set up drawn from the insights of the observations. As learning and potential failure lie at the heart of every experiment,

learning and action cycles are needed to advance and refine insights. After completion of the experimentation phase, the collected data can be used for a final evaluation.

Ongoing monitoring and the facilitation of participation and engagement

Experiments in real-world labs and related approaches regularly involve societal actors from various realms. The monitoring activities of an experiment also take place to a large extent in contact with those involved and affected. Thus, ongoing monitoring should be embedded in a communication strategy with participants and other stakeholders. Here, the various monitoring activities, such as questionnaires, interviews, or (small) group discussions, not only make it possible to collect data and keep track of the level of support for the experiment. They also make it possible to convey information to the participants and to involve them in the (re)design of the experiment. Visible consideration of the monitoring results obtained in this way can strengthen the confidence and motivation of those involved. Monitoring directly impacts central aspects in the dimensions of input and processes of the experiment (see figure 1). A lack of uptake of the feedback sometimes has the opposite effect: data collection also raises expectations. Thus, the monitoring activities can serve further purposes in the experiment and real-world lab. Where do synergies arise from monitoring and shaping participation? Where should these activities remain separate? What kind of participation is targeted in monitoring, and for what purpose?

A classic way of differentiating the degree of participation according to its intensity and objective is the participation ladder, reaching from 1) unilateral information giving to 2) simple, one-side consultation, 3) mutual exchange, 4) cooperation, and 5) empowerment including the permanent transfer of responsibility to the participants (Arnstein 1969; for a recent development see Stauffacher, M.; Krütli, P.; Flüeler, T.; Scholz 2012). The level of participation should be based on the extent of the desired effects. A rule of thumb: the greater, more comprehensive, and longer-term the effects are to be, the more intensive, diverse, and early participation should be (Newig et al., 2019). These considerations are concretized below.

A practical example of monitoring and the facilitation of engagement

In an international research and practice project, eight universities collaborated, sharing a focus on developing sustainability solutions[iii]. The project took place in 2017 and was part of the Global Consortium for Sustainability Outcomes initiative. As one element of the project, three universities in the United States, the United Kingdom, and Ireland planned to remove the non-essential hot water supply in the sanitary areas of selected buildings. The main aim was to reduce the massive energy consumption and greenhouse gas emissions associated with providing hot water. If successful, this activity could be scaled-up by implementing it in other buildings and other universities.

Both technical and social interventions were utilized. Technical intervention included “the installation of flow meters and temporary turn-off of hot water supply depending on the local supply system (e.g., steam, boilers)” and the provision of “essential hot water […] by installation of point-source heaters” (GCSO 2018:3). Social intervention included “various social engagement measures (e.g., surveys, focus group discussions) applied to different buildings for comparative purposes “(GCSO 2018:3). Overall, the project team considered it

essential to involve and secure the support of the building users, especially to avoid contrary user behavior (for instance, people not washing their hands). More intensive participation also seemed possible, not only to improve the information base of the project but also to catalyze wider changes in resource use.

As authors, we were a part of the larger project team and supported the monitoring and evaluation efforts as well as those aiming to facilitate participant engagement. We had to consider the available resources to conceptualize and implement these measures. The timetable for implementing the experiments was rather tight at four months. In addition, the personnel resource base was limited. University partners had their foremost expertise in the technical-economic field. "Since the project began with a specific technological intervention (the turn-off of non-essential hot water) planned for a set timeframe, this pre-determined technology-centered approach considerably reduced the possibilities for more meaningful engagement of stakeholders in co-designing the method(s) of intervention to be taken in the project" (GCSO 2017:5). Accordingly, the project partners decided for low to medium intensity participation as the most feasible appropriate (levels 2-3 in table 2). Accordingly, the measures aimed at reducing contrary user behavior (e.g., due to ignorance), adapting the experiments to actual user behavior and needs and increasing the acceptance of the experiments. We drafted specific engagement plans for each location, "building on the aims of the overall project in general and aims of each implementing university in particular" (GCSO 2017: 5).

On the monitoring side, we used the generic monitoring and evaluation scheme presented above (see figure 1, Luederitz et al., 2017) to conceptualize the dimensions and features of the assessment. Again, considering resource constraints, we decided on a lean selection of features to be monitored and straightforward measurement tools (see table 2 for an overview). We did structured interviews with key informants and recommended an online baseline survey to gather data on different aspects of the input dimension, such as awareness, engagement, and trust. Further, we recommended ongoing surveys and individual interviews and facilitated group discussions to capture, at least in parts, process aspects and emerging outcomes. Ongoing consumption data collection provided by the university's technical staff served to measure water and energy consumption and thus detect direct, real-world changes. The combined monitoring and participation measures were specified in a schedule in terms of timing, responsible persons, and target groups (recall table 1 for a generic example).

Monitoring and comparing the engagement facilitation measures led to the following key insights (see GCSO 2017: 7 ff for details):

- "Engagement intensity level needs to be planned well in advance to adapt the project to local conditions; this is particularly true for more intensive engagement formats […].
- Engagement results need to be fed back into the project design. The combination of engagement measures and elements of formative, ongoing evaluation appears promising/suitable for effective and iterative project management.
- Engagement should start as early as possible […]
- Key stakeholders can vary widely between buildings […, yet] there appears to be a number of common key stakeholder groups in all locations (e.g., owners, general management, users, and health and safety officers).

- Seemingly small interventions may raise considerable resistance [for instance, if associated with pre-existing concerns or if there is no] compelling argument for the purpose or need of the intervention. This reinforces the need for engagement early-on, that is targeted and meaningful."

Conclusion

Monitoring and evaluation are considered essential to transforming a random intervention into an experiment capable of generating scientific insights *and* societal impact. Societal engagement, however, is key to the actual implementation of a participatory experiment and a core ingredient of every monitoring and evaluation attempt. With the elements presented above, we aim to support the practical monitoring, adaptation, and evaluation of participatory experiments and to develop possibilities to fruitfully combine monitoring, evaluation, and the facilitation of engagement. From our own experience, we can say that this is not easy nor necessarily successful. Nonetheless, considering the co-creative and transformative appeal of participatory experiments and related labs, it is a worthwhile and essential endeavor. Further practical and scientific work is required to elaborate on how to raise synergies amongst the two essentials of participatory experiments.

Table 2: Concretization of participation levels for a socio-technical experiment at university campus: objectives, methods, and feasibility.

Level of participation	**Description**	**Concrete objective**	**Exemplary methods for shaping participation**	**Feasibility**
5. Empowerment	Transfer of decision-making power.	Transformation of energy use behavior inside and outside the university.	Identify key stakeholders, support their own project ideas, enable exchange and learning, mediate conflicts, *supervision/ formative evaluation.*	Not feasible/ appropriate due to limited project time, budget constraints, and pre-decided technical intervention.
4. Collaboration	Equal partnership with stakeholders.	Involving stakeholders in the design of the intervention from the beginning, sharing responsibility for outcomes and continuation of the intervention.	Co-design: regular meetings, collaboration from design to implementation to evaluation of the intervention, joint teaching.	Limited feasibility, due to short overall project duration and pre-decided intervention. Inclusion partly appropriate and feasible in teaching & activity of student groups.
3. Cooperation	Mutual exchange, decision-making power remains with researchers.	• Timely inclusion of stakeholder interests to improve the design of the intervention. • Identify and reduce side effects and risks. • Build awareness.	• *Informal group interview*s with different stakeholder groups to prepare engagement and understand the situation on the ground. • *Analysis of* possible risks and side	Probably the most appropriate level to make maximum use of a technical intervention that is already decided; feasible to a certain degree due to limited time and budget to realize engagement.

		• Reduce side effects due to ignorance.	effects of the intervention. • *Regular presence on site.* • Accessible information on site and online. • Email contacts and website maintenance.	
2. Consultation	One-sided stakeholder consultation.	• Gain insights into resource use. • Identify possible side effects. • Baseline assessment / baseline survey.	*Structured interviews* and *baseline survey* before intervention (prepares inclusion and collaboration).	High feasibility: any form of interviews for the baseline-assessment will be seen as consultation by the stakeholders.
1. Information	One-sided information of stakeholders.	• Informing about the meaning and aim of the intervention. • Reduce opposition and contrary behavior.	Signs, information boards, info emails, short presentation, *circular emails, personal talks* (depending on existing communication channels).	Very high feasibility: providing additional information, e.g., regarding hygiene.
0. No engagement	---	Simple technical intervention (e.g., turning off hot water).	None.	Already scheduled.

Acknowledgments

This is a translated, strongly modified version of Schäpke, Niko; Beecroft, Richard (in Print): *Laufendes Monitoring, allfällige Anpassungen und die Förderung von Engagement in partizipativen Experimenten*. In: Selma L'Orange Seigo, Matthias Probst, Michael Stauffacher, Eveline Lobsiger, Yann Blumer (Hrsg.): Interventionen in Reallaboren: Ein Handbuch für die Praxis. ETH Zurich

Figure 1 is reprinted from: *Learning through evaluation – A tentative evaluative scheme for sustainability transition experiments*, C. Luederitz, N. Schäpke, et al. Arnim Wiek, Daniel J. Lang, Matthias Bergmann, Joannette J. Bos, Sarah Burch, Anna Davies, James Evans, Ariane König, Megan A. Farrelly, Nigel Forrest, Niki Frantzeskaki, Robert B. Gibson, Braden Kay, Derk Loorbach, Kes McCormick, Oliver Parodi, Felix Rauschmayer, Uwe Schneidewind, Michael Stauffacher, Franziska Stelzer, Gregory Trencher, Johannes Venjakob, Philip J. Vergragt, Henrik von Wehrden, Frances R. Westley, Journal of Cleaner Production, Volume 169, 2017, Pages 61-76, https://doi.org/10.1016/j.jclepro.2016.09.005. Copyright (2017), with permission from Elsevier.

This article benefitted from earlier considerations on guiding transition experiment evaluation, as they were part of C. Luederitz, D.J. Lang, N. Schäpke, A. Wiek (unpublished): Evaluation of Sustainability Transition Experiments: Initial Considerations on Guidelines for Application of the Tentative Evaluation Scheme. Working paper. Version 2016. Leuphana University Lüneburg.

References

Arnstein, S. R. 1969. "A Ladder of Citizen Participation." *Journal of the American Institute of Planners* 35 (4): 216–24.

Beecroft, R., H. Trenks, R. Rhodius, C. Benighaus, and O. Parodi. 2018. "Reallabore Als Rahmen Transformativer Und Transdisziplinärer Forschung: Ziele Und Designprinzipien." In *Transdisziplinär Und Transformativ Forschen. Eine Methodensammlung.*, edited by Rico Defila and Antonietta Di Giulio, 75–100. Wiesbaden: Springer VS. https://doi.org/doi:10.1007/978-3-658-21530-9_4.

Caniglia, G., C. Luederitz, T. von Wirth, I. Fazey, B. Martín-López, K. Hondrila, A. König, et al. 2020. "A Pluralistic and Integrated Approach to Action-Oriented Knowledge for Sustainability." *Nature Sustainability*. https://doi.org/10.1038/s41893-020-00616-z.

Caniglia, G., N. Schäpke, D.J. Lang, D.J. Abson, C. Luederitz, A. Wiek, M.D. Laubichler, F. Gralla, and H. von Wehrden. 2017. "Experiments and Evidence in Sustainability Science: A Typology." *Journal of Cleaner Production* 169. https://doi.org/10.1016/j.jclepro.2017.05.164.

Fazey, Ioan, Niko Schäpke, Guido Caniglia, James Patterson, Johan Hultman, Barbara van Mierlo, Filippa Säwe, et al. 2018. "Ten Essentials for Action-Oriented and Second Order Energy Transitions, Transformations and Climate Change Research." *Energy Research & Social Science* 40 (April 2017): 54–70. https://doi.org/10.1016/j.erss.2017.11.026.

Global Consortium for Sustainability Outcomes (GCSO). (2018). Final Report of the Urban-Rural On-Campus Solutions: Sustainability Accelerator Program, "Hot Water Project". Online available (6.11.2022): https://gfl.wp.rtd.asu.edu/sustainabilityoutcomes/wp-content/uploads/sites/3/2018/05/GCSO-UROC-Final-Report.pdf

Lam, David P. M., Berta Martín-López, Arnim Wiek, Elena M. Bennett, Niki Frantzeskaki, Andra I. Horcea-Milcu, and Daniel J. Lang. 2020. "Scaling the Impact of Sustainability Initiatives: A Typology of Amplification Processes." *Urban Transformations* 2 (1): 3. https://doi.org/10.1186/s42854-020-00007-9.

Luederitz, C., N. Schäpke, A. Wiek, D.J. Lang, M. Bergmann, J.J. Bos, S. Burch, et al. 2017. "Learning through Evaluation – A Tentative Evaluative Scheme for Sustainability Transition Experiments." *Journal of Cleaner Production* 169. https://doi.org/10.1016/j.jclepro.2016.09.005.

McCrory, Gavin, Niko Schäpke, Johan Holmén, and John Holmberg. 2020. "Sustainability-Oriented Labs in Real-World Contexts: An Exploratory Review." *Journal of Cleaner Production*. Elsevier Ltd. https://doi.org/10.1016/j.jclepro.2020.123202.

Newig, Jens, J. Lang, Daniel, Stephanie Jahn, Judith Kahle, and Matthias; Bergmann. 2019. "Linking Modes of Research to Their Scientific and Societal Outcomes. Evidence from 81

Sustainability-Oriented Research Projects." *Environmental Science and Policy* 101: 147–55. https://doi.org/10.1016/j.envsci.2019.08.008.

Norström, Albert V., Christopher Cvitanovic, Marie F. Löf, Simon West, Carina Wyborn, Patricia Balvanera, Angela T. Bednarek, et al. 2020. "Principles for Knowledge Co-Production in Sustainability Research." *Nature Sustainability* 3 (3): 182–90. https://doi.org/10.1038/s41893-019-0448-2.

Schäpke, N., F. Stelzer, G. Caniglia, M. Bergmann, M. Wanner, M. Singer-Brodowski, D. Loorbach, P. Olsson, C. Baedeker, and D.J. Lang. 2018. "Jointly Experimenting for Transformation?: Shaping Real-World Laboratories by Comparing Them." *GAIA* 27. https://doi.org/10.14512/gaia.27.S1.16.

Stauffacher, M.; Krütli, P.; Flüeler, T.; Scholz, R. W. 2012. "Learning from the Transdisciplinary Case Study Approach: A Functional-Dynamic Approach to Collaboration among Diverse Actors in Applied Energy Settings." In *Tackling Long-Term Global Energy Problems*, edited by J. Spreng, D.; Flüeler, T.; Goldblatt, D.; Minsch, 227–45. Dordrecht: Springer.

Williams, Stephen, and John Robinson. 2020. "Measuring Sustainability: An Evaluation Framework for Sustainability Transition Experiments." *Environmental Science & Policy* 103 (October 2019): 58–66. https://doi.org/10.1016/j.envsci.2019.10.012.

---

[i] The addition of "participatory" distinguishes experiments in e.g., real-world labs from conventional experiments with a high degree of control over experiment and context, for example in clinical laboratories (see Caniglia et al. 2017).

[ii] An evaluation that follows an experiment and conclusively assesses it can also be structured in a similar way.

[iii] The example was part of a cooperation project of the Global Consortium for Sustainability Outcomes. The following universities cooperated: Arizona State University, Dublin City University, King's College London, Hong Kong University of Science and Technology, Instituto Tecnológico y de Estudios Superiores de Monterrey, Universidad Nacional Autónoma de México, Karlsruhe Institute of Technology, Leuphana University Lüneburg. See https://sustainabilityoutcomes.org/off-campus-solutions-sustainability-accelerator-program/