fteval JOURNAL
for Research and
Technology Policy Evaluation

July 2025, Vol. 57, pp. e11, 1-12
DOI: 10.22163/fteval.2025.703

EVALUATING "LEARNING AND EXPERIMENTAL SPACES"

WHEN A TRADITIONAL APPROACH REACHES ITS LIMITS

CHRISTINA SCHUH, DANIEL SCHWERTFEGER AND SONJA FRINGES
DOI: 10.22163/FTEVAL.2025.703

ABSTRACT

The paper reflects on the evaluation of a case study that seeks innovative solutions for digital transformation. In particular, it discusses the tension between traditional evaluation approaches and new perspectives on the funding process in general, the used methods and the changing role of the evaluator. The subject of the evaluation is the funding guideline of the German Federal Ministry of Labour and Social Affairs: "*Sustainable companies and administrations in digital change*". The Learning and experimental spaces (LES) funding instrument supports small and medium-sized enterprises in developing innovative, tailor-made and consensual solutions for employees and companies in the digital transformation. These solutions must also be supported by social partnerships. The several LES should be of a fundamentally exemplary nature and transfer to further innovative solutions. Funding was provided for 17 LES in the first funding round (starting 2018) and a further 11 LES with focus on the use of artificial intelligence (AI) in the second funding round (starting 2020). The duration of each funded project was about three years. Each individual project was evaluated externally and additionally subjected to an overall evaluation by the Federal University of Applied Administrative Sciences (HS Bund), which included all projects of the respective funding round. The paper first provides a theoretical framework for the background of the funding guideline. The second part presents the LES funding guideline and its evaluation, including some project examples. Then we describe and discuss the evaluation process, using five factors to compare traditional evaluation approaches and new perspectives on the funding process and the changing role of the evaluator before we end with our conclusion.

Keywords: Learning and experimental spaces, small and medium-sized enterprises, linking digital transformation and social innovation, consensual solutions, developmental evaluation

1. INTRODUCTION

For several decades, the need to adapt innovation policy as well as the regarding evaluation paradigms and practices to the changed societal and environmental problems has been in the focus of attention. Since the mid-20th century, at least two paradigm shifts took place in innovation policy (IP) and along with corresponding changes in methods of assessment and evaluation (Schot & Steinmueller, 2018; Rohracher, Coenen and Kordas, 2022). Early IPs focused on innovation mainly for economic growth, prosperity and mass production, whereas after the first shift, IPs started dealing with international competition and the link between discovery and application. After another important paradigm shift, the focus is now on transformative change, that is, addressing major global and societal challenges such as the Sustainable Development Goals. Thus, for some time now, we have increasingly been confronted with wicked problems (Reale, 2021) which do not allow for a single solution (if there is even one) or predefined solutions (Rittel & Webber, 1973). As Schot and Steinmueller (2018) conclude, the model of innovation must be experimental in this paradigm. Despite this insight, a gap often remains between the claim of transformative IPs and their actual implementation, including the practices of their evaluation (Rohracher, Coenen & Kordas, 2022).

In this practice report, we describe our experiences that the implementation of interventions and funding as well as the expectations of (some of the) stakeholders can make it difficult to apply flexible evaluation methods.
The subject of our report is a funding guideline from the German Federal Ministry of Labour and Social Affairs entitled “Sustainable companies and administrations in digital change”. A total of 28 projects were funded in this program, which were designated *Learning and experimental spaces* (LES), as they focus on learning experiences rather than products. The application of flexible evaluation methods is essential to address a key characteristic of wicked problems, namely, that each is a one-shot problem (Rittel & Webber, 1973). To gain knowledge of promising innovation pathways, one needs to gather the experiences of different actors with different perspectives (Schot & Steinmueller, 2018). This is far from trivial in the case (presented below), where

an overarching evaluation was conducted across several distinct and highly diverse projects.

An example of an approach, that could meet the requirements of the transformative paradigm is, developmental evaluation (Patton, 2010), which allows flexible application of methods and an active role of the evaluator. It takes a systems- and innovation-oriented approach and focuses on adapting interventions to changing contexts, target groups, or emerging needs. Therefore, developmental evaluation involves flexible designs, flexible relationships, flexible budgeting and flexible reporting. Accountability in this approach is extended to accountability for learning, development and adaption, and evaluators are supposed to be part of the evaluation team (Patton, 2015).

2. OVERALL EVALUATION OF THE LES FUNDING PROGRAM

As introduced above, in 2017 the German Federal Ministry of Labour and Social Affairs announced a funding guideline entitled "Sustainable companies and administrations in digital change"[1], for which consortia of small and medium-sized companies and application-oriented research institutions, including universities, applied for project funding. A central concern of the funding is, to implement the connection between technological and economic change processes and social innovation within the framework of strong employee and company participation. The institutional framework for this is provided by the BMAS's *New Quality of Work Initiative* (INQA), which is based on social partnership and designed to promote innovative solutions. The aim was to promote innovative, tailored and consensual solutions for employees and companies in the digital transformation, supported by the social partners, e. g. work councils or trade unions. Company learning and experimental spaces should be fundamentally, exemplary in character and contribute to further innovative solutions. The objects of funding were so-called *Learning and experimental spaces (LES)* – interventions for which failure was allowed, and measures could experimentally be tested out. The 17 plus – in a second AI-oriented funding round[2] – 11 projects were all evaluated individually by

1 See BAnz AT 14.08.2017 B2 at https://www.bundesanzeiger.de (last visited 23.05.2025).

2 See BAnz AT 20.01.2021 B2 at https://www.bundesanzeiger.de (last visited 23.05.2025).

separate independent institutions. Our team of the Federal University of Applied Administrative Sciences was commissioned to conduct an overall evaluation of the guideline.

The funding instrument of the learning and experimental space is considered particularly suitable for finding such strategies due to its open-ended approach, which enables operational learning and experimentation with new forms of work as iterative processes. An essential feature of this funding logic is that 'failure' is permitted, as setbacks are seen as part of the learning process. This funding logic is intended to free projects from the restrictive (implicit or explicit) expectation that they must present 'working' innovations or products at the end of the funding period. Project results can therefore also be learning outcomes that arise from trying out new technologies, methods or forms of work.

All funded project interventions aimed to promote digitalization in and medium-sized enterprises. For example, bus drivers in Leipzig were equipped with tablets to enable better communication with each other. In the care sector, speech recognition software was tested to facilitate documentation. In another project, exoskeletons were used by companies specializing in the renovation of bathrooms to support heavy work.

In order to better understand the overarching objectives of the funding guideline, a few more detailed insights into selected projects are provided here: One exemplary project from the 17 projects in the first funding round was entitled *AgilKom*. This project brings together stakeholders from the local administration (the city of Essen and the administrative district of Soest), the United Services Union (ver.di) and the German County Association. The specific intention of the *AgilKom* project was to implement innovative and agile processes in the administration on a technological and organizational level. And contrary to the usual top-down-logic, such solutions were developed in so called *Innovation-Labs*, in which employees work together across hierarchies and disciplines. The overarching goal in the end was to improve flexibility, efficiency as well as closeness to citizens of the administration. Regarding the impact level, the project aspires to gain insights into the transferability of central principles of agile organization to the public sector. Another project was *Handwerksgeselle 4.0*, which was about the development of technological assistance systems in the sanitary, heating and air conditioning industry. Accordingly, stakeholders from both, the development sector and the application sector were working together in this LES with the aim of solving industry-typical problems (e.g. shortage of recruits, demographic change

and competitive pressure). An exemplary project from the 11 projects in the second, AI-oriented funding round is the *KIDD* project. While the project primarily involved small and medium-sized enterprises, major corporations also contributed by collaborating on various topics (e.g. sales, evaluation of services, personnel management). With a focus on the application of digital systems in the work context, the KIDD project aimed to promote diversity among employees. This is achieved by developing standardized processes and criteria for these systems – especially in the application of AI.

As overall evaluators, our mandate included, on the one hand, compiling a synopsis of the findings from the individual evaluations across the 17 and 11 projects. During the clarification of the mandate, however, we also realized that the ministry wanted to know how well the new funding logic and the LES as funding instrument were working. This led us to expand the mandate to include this point on the other hand.

In concrete terms, the daily work of the overall evaluation consisted largely of networking and maintaining contact with the individual project evaluations. Data and findings from the individual experimental spaces, usually gathered at the beginning, throughout and at the end of the funding period, were incorporated into a survey instrument, developed specifically for the overall evaluation.

We drew on the approach of developmental evaluation (Patton, 2010), without claiming fidelity to its pure form or essential principles (Patton, 2016). Developmental Evaluation focuses on providing real-time feedback to support decision-making in complex, evolving environments. It prioritizes learning over accountability, helping stakeholders adapt strategies as programs unfold. Developmental Evaluation is highly collaborative, with evaluators working closely with program staff and stakeholders to co-create solutions. Unlike traditional evaluations this approach is flexible and emergent, adapting its methods and goals as the program evolves. The approach is designed for innovative or change-driven initiatives, emphasizing systems thinking and the continuous refinement of processes. We see a connection between the learning and experimental spaces and Patton's (2015) principles of utilization-focus, developmental evaluation and co-creation principles that both, the project evaluators and our team aimed to uphold. This involved flexibility of the methods as well as dynamic designs (Patton, 2015) that adapted to the changing timelines and contents of the different projects. The reference to Patton became apparent during the evaluation process. When planning an overall evaluation, the benefits were not yet clear to us, which illustrates our adaptive approach.

Thus, based on the insights gained during the evaluation process, we applied an adaptive multi-method approach (see Figure 1). Besides a document analysis (see first column), the evaluation included several components. First, on the qualitative side (second column), we conducted guided, individual face-to-face interviews with representatives from all stakeholder groups, involved in the funding instrument LES. We also held focus groups with volunteer evaluators and responsible representatives from the ministry. In addition, we asked for written reflections from the project managers on the course of the project and on the funding instrument. Furthermore, in the third column: (longitudinal synopsis), we used an instrument to collect what we called descriptive fields. These were based on a framework and included recommended indicator categories such as learning and application, improved working conditions, sustainability and future viability, communication, participation and organization. We also conducted a standardized online survey with project managers (see last column).

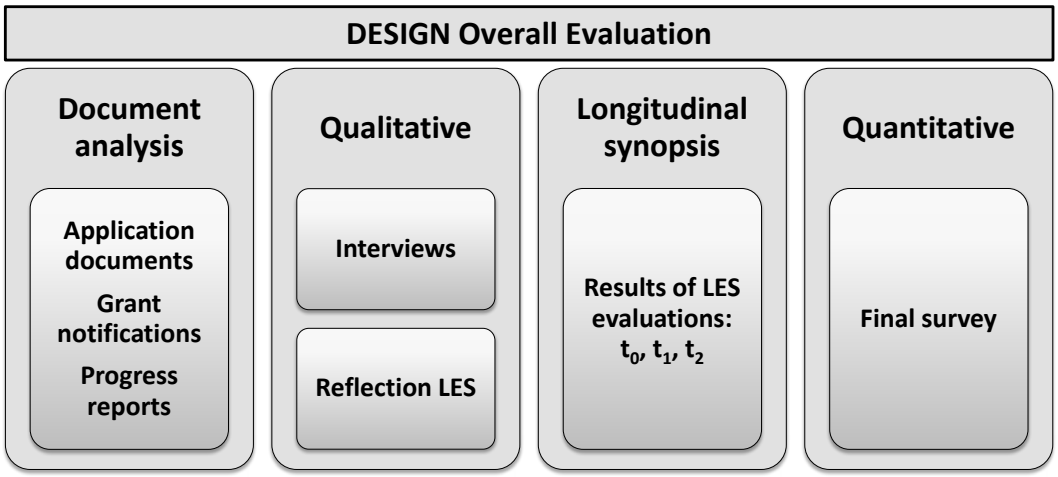

Figure 1: Design Overall Evaluation

3. REFLECTION ON THE EVALUATION PROCESS

In the following, the evaluation process is examined from the five perspectives: the financing logic (asking: How well did the linear path function?), the context in which the measures of the individual projects were implemented, the specific interventions, the methods applied and the role of the evaluators.

FINANCING LOGIC

The idea is that experimental spaces, in contrast to linear project funding, should make it possible to make mistakes and to allow failure during the research project. We found this idea has been well received by the project participants, as it was mentioned in both interviews and focus groups as a special feature of the funding instrument. However, the more interesting question is, to what extent these possibilities have been implemented during the project. We found indications that there is still room for improvement in this area. On the one hand, only a few projects documented processes of failure, on the other hand, there were also only few instances where projects deviated from their original goals or budget planning and made corresponding adjustments. One possible reason for this might be that the additional administrative effort involved in the projects (iterative cycles) was not considered. The challenge of evaluating non-linear funding logics, aimed to addressing wicked problems is described in the literature as lying primarily in several areas of tension: the attribution of outcomes to specific interventions, limited funding periods versus impact measurement, processes versus outcomes and reflexive learning versus external control (Rittel & Webber, 1973; Rohracher, Coenen and Kordas, 2022). In our overall evaluation, two-thirds of the project managers, surveyed in the final round, judged the three-year funding period as too short to effectively measure impact. This despite the fact that the funding guideline, as the name suggests, explicitly focuses learning. From an academic point of view, as represented by the technical support provided by the departmental research institution BAuA (Federal Institute for Occupational Safety and Health), the question of attributing intervention success and identifying causal effects remains largely unresolved. Instead, it is primarily anecdotal knowledge that has been generated, which could potentially be applied successfully to other areas. Short-term solutions to long-term problems cannot be implemented and therefore not measured. The consequence for the evaluation process is that it adopts a process-oriented, formative view instead of an ex-post, summative one. From an epistemological term, it shifts away from a predominantly positivist approach toward a more interpretive and constructivist one.

CONTEXT OF INTERVENTION

Solving complex problems with the help of project funding, requires accurate understanding of wicked problems. This includes an understanding of the possibilities of funding logics and the acceptance that compromises, made in the attribution of success and in the transfer of results. The 28 LES projects shared the characteristic of operating in highly dynamic and complex environments, especially when viewed in the context of the funding guideline. As mentioned above, the heterogeneity of the projects made it difficult to summarize their results, an issue stemming from the diversity of funding guidelines and subsequent results. On the other hand, this heterogeneity enables transferability to other domains at a sufficiently high level of abstraction, which is reflected in some of our recommendations. For example, in several projects we were able to show that expectations of the intervention were often too high. Accordingly, managing expectations realistically helps to more accurately reflect the potential impact of the measures. This could be achieved by specifying the technological focus in the funding guideline from the outset. Another important factor for all LES projects conducted over the past four years is COVID-19. The pandemic has led to a new way of dealing with uncertainty and unforeseeable situations. In response to this, we mainly received the message that there was a kind of *COVID-Boost,* in terms of the acceptance of digitalization measures. In summary, it is not only difficult to summarize the results of heterogeneous projects, but also to compare them meaningfully. This problem is exacerbated when these projects are implemented in complex and uncertain environments.

SPECIFIC INTERVENTIONS

The solutions within the individual projects were co-developed with the social partners and had not been predetermined at the project's outset. However, it remains unclear to what extent these solutions were further developed, or even discarded, over the course of the project. After all, this rather gradual development of interventions and solutions would have been entirely in line with the concept of the funding guideline. We gained the impression that the available opportunities for experimentation had not been fully utilized. However, it does not matter that the opportunities, created by the new intervention logic, will be utilized. Even if there is a will to make funding options more flexible to meet the need of transformational IPs, there must also be the courage to implement them (Rohracher, Coenen and Kordas, 2022).

APPLIED METHODS

From a methodological perspective, a challenge arose in conducting the overall evaluation regarding the fields of indicators recommended by the technical support from BAuA. These indicators were also recommended to the individual projects at the start of the funding period, without our ability to influence which ones would be used. In other words, there was no fixed set of indicators available at the outset that we could offer to all individual projects for reporting purposes. Rather, it was clear to us from the outset that we would develop such an instrument together with the evaluators of the individual projects. In our view, reflection on applicable indicators is generally linked not only to the overarching funding objectives, but also to how these can be addressed in cooperation with the individual projects. However, bundling the partially qualitative information proved difficult. The needed effort increased further over the course of the funding period, especially with regard to the AI funding round. In conceptual terms, we were able to conduct a formative evaluation because we were part of the funding guideline from the beginning of the projects. The synoptic presentation of the evaluation results over three measurement points was developed using a coding system for the data from the first measurement (t_0). Three coders employed a bottom-up approach to develop and compare responses from three randomly selected documents. For questions with largely identical content across later survey dates, supplementary codes were assigned where necessary and subsequently discussed to reach a consensus. The resulting coding system served the basis for the content and synoptic analysis of the three survey dates. The extent to which our overall evaluation and the synopsis of individual results have contributed to addressing the major problems must be assessed by others. In any case, we were able to obtain valid results to improve a more recent call for proposals utilizing the instrument of *Learning and experimental spaces*.

ROLE OF THE EVALUATORS

A high degree of flexibility in applying and adapting the survey instruments also demands considerable flexibility in determining the required resources. Our understanding of our role as evaluators – and the BAuA similar understanding – was characterized by adherence to scientific standards, both in our methods and in the assessment of the individual projects. In the developmental evaluation approach, the evaluator is seen as part of the innovation team. This was the case neither in the individual projects nor in our overall evaluation. In the individual projects instead, the role of individual

evaluators was sometimes described as *critical friend* (Balthasar, 2012). The Ministry was ultimately responsible for the success of the funding guideline. However, our results also show that the project managers strongly identified with the interventions. On a subjective level, this represents an important additional factor for success. If the role of an evaluator is understood as a *critical friend* or even as *part of the innovation team*, like the developmental evaluation suggests, it can be stated that the role of the administrative actors is changing. In this way, the evaluators take on more of an advisory role and less of a hierarchical, top-down position.

4. CONCLUSION

Learning and experimental spaces are currently a widely used project funding instrument at municipal, state and federal level in Germany. Their evaluation requires a new perspective on the funding process, the role of the stakeholders and applied methods. Finally, based on our experiences, we would like to offer some conclusions for future overall evaluations. In any case, the primary challenge is to transition from anecdotal knowledge in very specific areas to identifying effective principles and patterns. A well-implemented intervention should only be extended to other environments if it has proven effective in a setting with a control group, something that is difficult to achieve in the real-world funding context. Focusing too much on transfer from the outset can also dilute the impact assessment. Furthermore, the *hunger for learning* – as opposed to the *fear of failure* – should be communicated and practiced from the outset. Although flexibility is required from both, the innovation team and its evaluators, evaluations of a certain scope cannot be conducted without some degree of top-down-control. It is not possible to collect valid and summable data in such a heterogeneous, agile and large-scale environment. The demands on evaluators' qualifications are once again increasing, in addition to qualitative and quantitative methodological knowledge, creativity and tolerance for ambiguity (skills we also put to test during in the pandemic) and strong social competencies are becoming increasingly important.

REFERENCES

Balthasar, A. (2012). Fremd- und Selbstevaluation kombinieren: der ‚Critical Friend Approach' als Option. In: Zeitschrift für Evaluation, 11(2), pp. 173-198.

Patton, M. Q. (2010). Developmental evaluation: Applying complexity concepts to enhance innovation and use. Guilford Press.

Patton, M. Q. (2015). State of the art and practice of developmental evaluation: Answers to common and recurring questions. In: Patton, M. Q., McKegg, K., & Wehipeihana, N. (Edts). Developmental evaluation exemplars: Principles in practice (Reprint). Guilford Press, pp. 1-24.

Patton, M. Q. (2016). What is essential in developmental evaluation? On integrity, fidelity, adultery, abstinence, impotence, long-term commitment, integrity, and sensitivity in implementing evaluation models. In: American Journal of Evaluation, 37(2), pp. 250-265.

Reale, F. (2021). Mission-oriented innovation policy and the challenge of urgency: lessons from Covid-19 and beyond. Technovation, 107, 102306.

Rittel, H. W., & Webber, M. M. (1973). Dilemmas in a general theory of planning. Policy sciences, 4(2), pp. 155-169.

Rohracher, H., Coenen, L., & Kordas, O. (2023). Mission incomplete: Layered practices of monitoring and evaluation in Swedish transformative innovation policy. In: Science and Public Policy, 50(2), pp. 336-349.

Schot, J., & Steinmueller, W. E. (2018). Three frames for innovation policy: R&D, systems of innovation and transformative change. In: Research policy, 47(9), pp. 1554-1567.

AUTHORS

CHRISTINA SCHUH
Federal University of Applied Administrative Sciences
Willy-Brandt-Straße 1
50321 Brühl, Germany
Email: christina.schuh@hsbund.de

DANIEL SCHWERTFEGER
Federal University of Applied Administrative Sciences
Willy-Brandt-Straße 1
50321 Brühl, Germany
Email: gesamtevaluation-experimentierraeume@hsbund.de

SONJA FRINGES
Federal University of Applied Administrative Sciences
Willy-Brandt-Straße 1
50321 Brühl, Germany
Email: gesamtevaluation-experimentierraeume@hsbund.de