
On the Evaluation of Futures Literacy Laboratories

Stefan Bergheim

Center for Societal Progress, and ZUKÜNFTE, Frankfurt am Main, Germany
stefan.bergheim@zukuenfte.net

Abstract

Futures Literacy Laboratories (FLLs) are a relatively new method for dealing with the future in complex adaptive systems. So far, little research has been done on the evaluation of the method's merits. This paper combines insights from evaluation research with results from ten interviews with experienced designers and facilitators of such laboratories. It shows that the acquisition of skills, the association among participants, the awareness of novelty, and the application of insights are recurring intentions of laboratories. Second, it deduces a set of feasible evaluation practices for such laboratories. And third, it highlights similarities between the design principles of Futures Literacy Laboratories and the principles of evaluation in complex adaptive systems.

Keywords

Futures Literacy, Evaluation, Complexity, Anticipation

1. Introduction

Futures work is rarely evaluated in a formal way, as Dator (2019, 292) bemoans. More solid evaluation could help the field prove the value of its work and increase its credibility with other academic disciplines, funding organizations, and users. In addition, the link between theoretical and practical work could be enriched by more evidence on which intervention works under what conditions (Gardner and Bishop, 2019). This paper was written for futurists who are interested in evaluating their work in complex adaptive systems. It investigates how evaluation is and can be done for one specific futures method that is relatively new to the field: Futures Literacy Laboratories.

Evaluation is defined as “the systematic assessment of the design, implementation, or results of an initiative [or intervention] for the purpose of learning or decision-making” (Canadian Evaluation Society, 2015). It investigates the value, merit, or worth of activities, and aims to identify what works, for whom, in what respects, in what contexts, and how. The purpose of the evaluation is to support the decision-making of funders and stakeholders, as well as the learning of designers of future interventions. As Chen and Hsu (2020, 103) put it, “evaluation is another form of learning” or of action learning as this paper will argue.

An intervention is defined in Merriam-Webster's dictionary as the act of interfering with the outcome or course of a process in a specific system. It is an activity designed to bring about change. The specific intervention in focus here will be described below: Futures Literacy Laboratories. The term “activ-

ity” could be used interchangeably, while “program” refers to longer-term activities.

Interventions usually have at least one intention or objective, something toward which the effort is directed: an aim, a goal, or an end. Interventions tend to lead to specific results or outcomes: Something is different afterward as a consequence or effect of it. Usually, there is a mapping between ex-ante intentions and ex-post results. Later, these results or outcomes may lead to an impact, although other factors are likely to influence this longer-term change as well.

An important challenge for evaluating futures work is that it is often done in complex adaptive systems, or more precisely: complex emergent anticipatory adaptive systems. These are well-researched since the work of Kauffman (1995), Cilliers (1998), and others: Multiple interacting and interdependent elements process information, adapt to impulses, and self-organize in often unpredictable ways. While there often are path dependencies, there is also emergent, discontinuous, and often non-linear change that is not knowable in advance. Using the terminology of Eoyang and Berkas (1999) these systems are dynamic, massively entangled, scale-independent, transformative, and emergent.

2. The research approach

The research process that led to this paper combines and connects two sources of knowledge to get a sense of what is feasible and what is done in practice in evaluating Futures Literacy Laboratories. Following a presentation of the method of Futures Literacy Laboratories in the next section, the paper provides an overview of approaches and principles of evaluation that were developed outside of futures studies, but which are helpful for the first part of the research project: Is it possible to do an evaluation in complex systems?

The essence of developmental evaluation and its relatives such as applied systemic program evaluation, collective impact evaluation, and systemic action research is described. We will show that there is a close correspondence – not expected at the beginning of this research – between design principles of laboratories and of evaluation in complex systems.

The second source of knowledge and of data created in this project are the insights from ten designers and facilitators of Futures Literacy Laboratories with many years of experience. The author is indebted to the following individuals for their valuable insights but bears full responsibility for this text. In alphabetical order including their affiliation at the time of the interview: Irianna Lianaki-Dedouli at PRAXI Network; John A. Sweeney at Westminster International University; Kwamou Eva Feukeu at UNESCO; Lilly Herde, Mirjalisa Walz and Sandra Schwarzman at Osnabrück University of Applied Sciences; Lydia Garrido at the South American Institute for Resilience and Sustainability Studies; Nick Balcom Raleigh at the Finland Futures Research Centre; Petra Cremers at Hanze University of Applied Sciences; Riel Miller at UNESCO; Roumiana Gotseva at the Center for Strategic Foresight; and Shermon Cruz at the Center for Engaged Foresight.

Structured interviews of between one- and two-hours length were held via videoconferencing in September and October 2021. This data was generated to provide insights on three further research questions: What is being evaluated in the context of FLLs? How is the evaluation done in practice? And, importantly, does this practice correspond to the principle of evaluation in complex systems?

After the interviews, all answers were added to a database. To ensure anonymity and possible use of the database by other researchers, the interview partners do not appear by name. Using the social science method of qualitative content analysis, each statement was linked to the interviewee, and to the question that led to it, and one keyword per statement was identified. These were mapped into categories that were created during the research by a combination of deductive reasoning from other evaluation research and inductive reasoning from the keywords. Comments and observations by the researcher were added to the statements to create a basis for the overall analysis later.

3. The intervention: Futures Literacy Laboratories

This paper focuses on one specific futures intervention: Futures Literacy Laboratories. Larger projects such as faculty development programs (Kazemier et al. 2021) are outside the scope here. Since Futures Literacy Laboratories are a relatively new method for dealing with the future and because several of its design principles are similar to the principles of evaluation in complex systems, the method will be presented in some detail here. It has been developed in particular by Riel Miller (2018) over three decades in close collaboration with researchers such as Roberto Poli. More than one hundred Futures Literacy Laboratories have been run by UNESCO and others around the globe over the past ten years with a large spectrum of participants and on a wide variety of topics, both onsite and online.

Futures Literacy Laboratories are rooted in the theory of anticipation, developed by the biologist Robert Rosen (2012) and explored for use in futures work in particular by Roberto Poli (2019). The core insight is that all living beings have predictive models of themselves and of their environment: anticipatory models. This means that “a change of state in the present occurs as a function of some predicted future state” (Rosen 2012, 8), bringing the Aristotelean final cause back into science.

Anticipatory assumptions in the present are the basic building blocks of these models and therefore the core unit of investigation in the laboratories. The spreading of anticipatory reasoning is labeled the “anticipatory turn” in social sciences in general and in the futures field specifically (Cevolini 2016). Laboratories are linked to the competence of futures literacy, which refers to the ability to imagine the future for different reasons using different models and assumptions.

Futures Literacy Laboratories are an addition to the small number of methods that Aaltonen (2009) labeled “social complexity” methods. These help people learn, imagine, and be creative about the future and allow ambiguity and emergence. Aaltonen placed Causal Layered Analysis (Inayatullah 2004), participatory methods, and visioning in this group. Futures Literacy Laboratories had not been part of his 2009 analysis. In contrast, Aaltonen

grouped a large number of methods under the label of “engineering approaches”, which rely on the ability of managers, experts, or researchers to understand, design, and control the system from the outside and identify clear rules: environmental scanning, genius forecasting, text mining, roadmapping, scenarios, and futures wheels. Different methods used for different reasons in different settings require different ways of evaluation.

Most Futures Literacy Laboratories comprise four phases that build on each other to make visible and experiment with anticipatory models and assumptions. Phase 1, the “reveal”, works with the expectations and hopes that participants have about the topic under investigation. This phase combines elements of trend analysis, forecasting, and visioning, with the explicit purpose of revealing the anticipatory assumptions that led to those images. Phase 2, the “reframe”, is a crucial learning catalyst and gives participants an opportunity to imagine the future of the topic through a frame that is unfamiliar and distinct from the one they revealed in phase 1. Participants experiment with a different set of anticipatory assumptions.

In phase 3, the “rethink”, participants come back into the present and look again at the three futures they explored in phases 1 and 2. They search for new issues that emerged for them, for fissures that open into new ways of thinking, and for powerful new questions that might turn into quests for them later. Finally, in phase 4, the “acting” or next steps, participants select from the wide menu of choices they created in phase 3 and focus on concrete next steps.

Three major design principles apply to Futures Literacy Laboratories (Miller 2018, 36-40). They are outlined here to illustrate what happens in practice and to lay the ground for the comparison with principles of evaluation in complex adaptive systems: First, these laboratories are action learning or learning-by-doing activities. They engage a variety of participants and invite them to articulate their individual images of the future and the underlying assumptions. As they do so, participants learn about a variety of ways to use the future. Burns (2015) describes the main features of action learning as generating theoretical as well as practical knowledge about the situation; enhancing collegiality, collaboration, and involvement of participants who are actors in the situation; establishing an attitudinal stance of continual change, self-development, and growth.

The second design principle of Futures Literacy Laboratories is that of collective intelligence knowledge creation: Learning is more likely to occur if it is done together with others rather than just by oneself. Participants can inspire each other, discuss meaning together and be more creative. This allows them to experience different forms of sensing and making sense, of knowing and not knowing. They explore specific individual perspectives as well as detect possible general patterns. In practical configurations of such laboratories, this design principle of collective intelligence has consequences for the selection of participants, for the design of the arrival situation, the size of breakout groups, and especially for the tools of learning together that range from talking with each other via drawing and sculpting to role-playing.

Third, these laboratories use a catalyst to disrupt entrained thinking, move attention beyond the old boxes, and stimulate emergence with different anticipatory models and assumptions. This is done in the second phase of the laboratory, the reframe, with an alternative future intended to take partici-

pants beyond the frames of likely-unlikely or desirable-undesirable. There are many ways to do this reframing, such as using static or dynamic configurations or games or arts.

All Futures Literacy Laboratories rely on the same general design principles and consider the four phases of reveal, reframe, rethink, and act. In practice, and importantly for evaluation, every laboratory is different, depending on the specific local circumstances. Most laboratories are held in cooperation with a local host or “local champion” who has an interest in the intervention, specific local expertise, and a network of possible participants. Some laboratories last a few hours, others several days. Some laboratories involve a small number of participants, others involve hundreds. Some laboratories focus on the creativity-enhancing phase 2, while others only briefly touch on that phase. Similarly, some laboratories focus on identifying and developing new activities in phase 4, while others deliberately end with phase 3 and rely on the energy of participants to work with the ideas after the laboratory. While this versatility may explain the recent spread of the method, it makes it more difficult to identify evaluation approaches that are applicable in general.

4. Evaluation in complex adaptive systems

As described in the previous section, some futures methods are more appropriate for complicated systems (“engineering approaches”) and others for complex systems. A similar distinction applies to evaluation approaches: If the result or outcome of an intervention is clearly defined at the start and stable over time, then engineering methods are appropriate. Bloom’s Taxonomy and Kirkpatrick’s Four-Level Model appear appropriate here.

By contrast, if the outcomes are emerging over the course of the intervention, other methods are needed. These have seen rapid development over the past twenty years, drawing inspiration from the works of Michael Quinn Patton (1978, 2010, 2019). This is the approach appropriate for Futures Literacy Laboratories and will be presented in this section in more detail as it may not be that well-known among futurists – and to highlight the parallels to the design principles of Futures Literacy Laboratories.

Some general principles and features of evaluation in complex adaptive systems have been formulated that appear relevant for the evaluation of Futures Literacy Laboratories and will serve as a cross-check for the results from the expert interviews in the next section. For example, Eoyang and Berkas (1999) map a set of evaluation principles into the five characteristics of complex adaptive systems mentioned earlier:

1. **Dynamic:** Evaluation should initially try to capture an emerging model of causal relationships to establish a baseline and revise that frequently. It seems necessary to evaluate and revise the evaluation design as new information comes in. One should also try to learn from the noise in the system and not focus only on the dominant or expected outcomes.
2. **Entangled:** Multiple evaluation strategies, methods, sources, or time horizons should be used because it is often impossible to tell which factor will be of particular interest at what time. Evaluators should be

explicit about the language and meanings of evaluation while designing it, collecting data, and analyzing them.

3. Scale independent: Information about the evaluation process and the findings should be accessible to all stakeholders to allow feedback at different levels. This makes evaluation a part of the intervention.
4. Transformative: Seeing evaluation as part of the intervention, it is crucial to involve as many elements of the system as possible in the design and implementation of the evaluation. And it should focus – especially early in the intervention – on reinforcing feedback with issues to celebrate.
5. Emergent: Evaluation has to match the developmental stage of the systems. If the system is highly unstable because of the intervention, it may not provide much meaningful information to be evaluated. It is also important to track patterns of change over time.

Cabaj (2014) describes five rules for developmental evaluation that are appropriate for complex adaptive systems:

1. Evaluation should enable – rather than limit – learning. The focus should be on learning-by-doing, with experimentation, ongoing tracking of outcomes, honest examination of failures, feedback from peers, community-wide reflections, and rapid adjustment if needed.
2. Employ multiple designs for multiple users. This rule relates to the principle of entanglement mentioned earlier. The additional effort for this flexible evaluation should lead to more relevant, useful, and timely results.
3. Shared measurement only if necessary. In many contexts, a requirement for shared measures may slow down the overall process, limit thinking, or get in the way of action or experimentation. Cabaj argues for “simple and roughly right” measures that support thinking and action.
4. Seek out intended and unintended outcomes. Interventions in complex systems tend to generate ripple effects that were not part of the original intentions. Evaluation should try to capture those as well.
5. Seek out contribution rather than attribution to change. Since multiple factors and conditions are behind any change, evaluation should not try to attribute outcomes to a particular group or intervention. Instead, it should try to investigate the contribution.

The similarities between principles and rules for evaluation in complex systems and the design principles of Futures Literacy Laboratories are obvious: both use action learning and collective intelligence knowledge creation. Both are highly flexible depending on the specific local circumstances. Both are open regarding the exact outcomes and do not require consensus. And they refer to or use an intervention or irritation to create new data about the system they operate in. Burns (2014) adds the need for an iterative process to assess impact in real-time and adjust action if needed.

5. Interviews and findings

The guiding questions developed for the ten interviews in this research were combined into seven groups. The first three groups relate to the intentions of designers, local champions, and participants of the laboratories. Why did they offer, hold, or participate in laboratories? Question clusters four and five were about what was evaluated and how it was done. The last two sets of questions related to the actual insights from the evaluation and to the use of those insights.

1. Why do you offer Futures Literacy Laboratories? Is there anything specific you want to achieve with these labs? This set of questions targeted the objectives, goals, purpose, needs, and intentions from the perspective of the interviewee.
2. Why do local champions or partners want to hold a lab? What are their reasons? This set of questions was a first step to go beyond the perspective of the designers and to include the views of a stakeholder group that was not interviewed for this project.
3. Why do you think people participate? This question widened the perspective further to include participants.
4. Do you or others evaluate or measure the success of the laboratories? If yes, what aspects are being evaluated? This question related to the initial why-questions about the intentions behind laboratories.
5. How do you or the team evaluate laboratories? How are data captured, processed, and analyzed? This set of questions related to the more practical aspects of evaluation activity, and also provided an opportunity to inquire about the evaluation models used.
6. Which results did the evaluations find? This question provided an opportunity to report on the actual outcomes of the laboratories.
7. In which way were the results used and by whom? To be worth the effort, the results of the evaluation should be used by somebody. This question tried to find out who the users were.

5.1 The intentions behind Futures Literacy Laboratories

All interviewees had a lot to say about the opening question, about why they themselves offer Futures Literacy Laboratories, and what they want to achieve with this method. One-third of the database entries relate to this question. The dominant reason is their expectation that they will support the acquisition of the competencies or skills of participants to imagine the future: their futures literacy.

Some of these expectations were formulated in general terms and appear to be difficult to evaluate directly: to help people think about the future, to help them use their imagination, experiment, be themselves, self-reflect, improve resilience, become more aware of complexity, etc. Other statements were more specific and might be easier to track in an evaluation: understand multiple ways to think about the future, enhance the cognitive capacity for prospection, see the constraints of their thinking, show that not knowing is okay, and become less scared.

The second major area of focus for why interviewees offer FLLs was the awareness of novelty. The term “aha-moments” was mentioned by three interviewees. Others referred to the awareness and discovery of opportunities and power structures not seen before, with the expectation that participants “cannot unsee”. Related statements included “open participants to things they don’t know”, “understand their context in different ways” or “enable better-informed decisions”. One interviewee indicated the possibility to use laboratories to create new ideas for academic research projects.

The association with others during and after the laboratory was mentioned as well, in line with the design principle of collective intelligence knowledge creation. Keywords used by interviewees and clustered by the author into this category include co-creation, conversation, connections, engagement, fun, and empathy. This appears to be an area amenable to evaluation as well.

The application of the insights and skills was only mentioned by two interviewees and only in very general terms: “foster transformation” and “support a better world”. Possibly, the long and complex causal chains made interviewees less focused on the application.

The answers to the second set of questions about the intentions of local champions show some overlap but also differences in the intentions of the interviewees. The acquisition of a competence seemed to be important for local champions as well. They wanted to “expand their foresight practice” and “increase their futures capacity as a competitive advantage”. And they wanted to “learn a new method about the future”. Overlapping keywords between designers’ and hosts’ intentions include creativity and novelty. Some hosts apparently felt “stuck with their old ways” and wanted to “get outside their normal boxes”.

In general, local champions seemed to have a stronger focus on the application of the insights gained during the laboratory and of the competence acquired, which was not a major focus of the interviewees themselves. According to the interviewees, local champions wanted to apply the results in their specific context for strategic work, leadership, decision-making, and to support transformation.

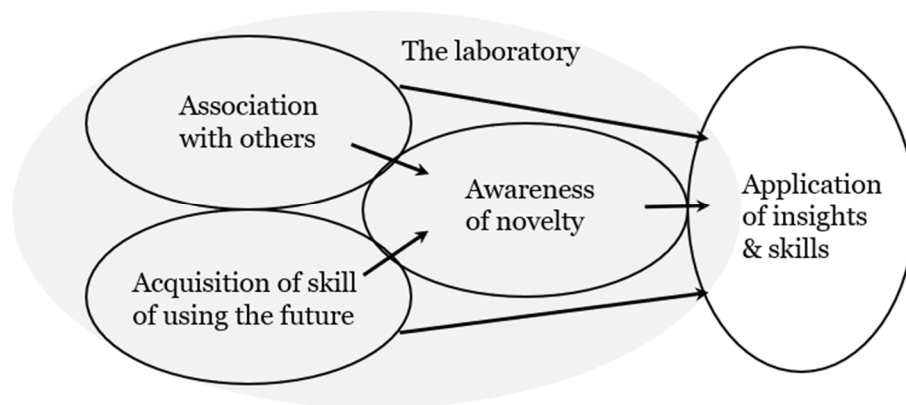
The third set of questions relating to the intentions of participants led to only a few answers. The reason could be that the designers and facilitators interviewed didn’t know much about participants because the local champions are the intermediaries to those. Among the few answers, the acquisition of higher futures competence was mentioned: learning a new method and focusing on the empowerment aspect. Networking and association also seemed to play a role for participants. Often, the interviewees’ impression of why participants took the time to join the laboratory was focused on the calling power of the local champion or of UNESCO.

The insights from the first three sets of questions about the intentions of laboratories were just presented using a structure that emerged during this research. It was inspired by the work of Bressers and Gerrits (2013) in combination with the qualitative content analysis of the interviews. The following model of the multiple and interrelated intentions and expected results of Futures Literacy Laboratories has four A-elements:

1. **Acquisition** of a skill, here the skill of using the future more broadly than before the laboratory, of better understanding anticipatory systems and processes.
2. **Association** with others is crucial in social complexity and in collective intelligence knowledge creation processes.
3. **Awareness** of novelty is an important element of a Futures Literacy Laboratory. Searching for, scanning, and identifying new knowledge, ideas, and questions on the topic is an important step before becoming aware of new possibilities for action. Awareness tends to come slightly later in the laboratory after the foundations for acquisition and association have been laid.
4. **Application** of the skills and of the insights usually comes even later in the process, sometimes after years when the conditions are suitable.

Inspired by Maurer (2021, 196) the following model illustrates these four elements and their rough sequencing. During the laboratory (grey area), the focus is initially on the individual's acquisition of the skill of using the future (e.g. separate probable and preferred futures, experiment with alternative futures) combined with the association with others in the collective intelligence process. Later during the laboratory, especially in phase 3, attention turns more to the awareness of novelty. The application and use of the new insights and of the new skills may start during phase 4 of the laboratory but happens mainly afterward in the participants' specific contexts, possibly in association with other participants.

Model of intentions of Futures Literacy Laboratories



Source: Stefan Bergheim

5.2 The what and how of evaluation

When the interviews turned to questions about what was evaluated and how that was done, several interviewees initially responded that they do not formally evaluate laboratories. However, later in the conversation, they all articulated that they had been doing an evaluation, possibly under another name or quite informally. Interviewees expressed several reservations about evaluation in complex systems: Few insights are gained if participants can reproduce after the laboratory what the facilitator told them about several times during the event. One interviewee mentioned the complexity of the

evaluation task given the multiple intentions of the different stakeholders. While this is a feature of any evaluation in complex systems, the overview of evaluation approaches earlier in this paper suggests that evaluation is feasible nevertheless. Another interviewee thought that evaluation was too problematic because the experience of the laboratory is specific for each individual person. While this is also common in complex systems, it should nevertheless be possible to get a sense of the general pattern of experiences.

Because of these reservations, the research database only contains entries from four interviewees on what they evaluate. As expected, those answers relate closely to the answers given on the reasons why they run laboratories: “participants strengthened their capacity to invent the unknown”, “collective learning took place”, “participants were excited and engaged”, “participants went outside their comfort zone and showed disorientation”, “participants understand how narrow their views on the future are”. One response related to the application: “People come back and ask for more”.

The fifth question on how interviewees or local champions did evaluation in practice generated a richness of methods and tools that appears surprising considering the general skepticism towards evaluation. It also showed a wide spectrum of actual evaluators ranging from participants via designers and facilitators to external evaluators.

- Some interviewees used surveys or questionnaires, one even online during the laboratory. One interviewee mentioned caution towards surveys because they can be disruptive, especially when used before the laboratory. This can be the case, but if evaluation is used as an integral part of an intervention (as suggested by the principles of evaluation outlined above), it can support the intended disruption or learning. Another interviewee mentioned that surveys appear inappropriate in emotive fields of novelty.
- Four interviewees mentioned the use of self-reflections by participants. They had asked them to identify the most meaningful moments during the lab, to describe their journey, to play with dixit-cards, to play emotions pantomime, or to check out using adjectives about their experiences. One critical issue could be whether those reflections mostly produce socially acceptable results and therefore might miss some of the true thinking and emotions.
- Three interviewees mentioned explicitly that they themselves observe participants’ reactions and engagement during the laboratory. Bodily expressions such as nodding heads, smiling or attentive listening can give a (highly subjective) impression about whether the design and facilitation are appropriate.
- Semi-structured interviews with participants were also mentioned, but mostly as options or plans. These seem to be promising vehicles to assess the application of new insights and competencies several months or years after the laboratory.
- Retrospectives among designers and facilitators as well as with the local champion were mentioned by half of the interviewees. In most cases, these retrospectives did not follow a specific structure and they were not recorded.

-
- Only one interviewee mentioned that an evaluation report was written after a laboratory.

By combining these practices with approaches mentioned in the evaluation literature, the following sequence of evaluative activities around Futures Literacy Laboratories – and potentially other futures methods in complex systems – emerged during this research. These elements reflect the approaches of practitioners and are generally in line with the principles for evaluation in complex systems described above. In concrete cases, specific elements would have to be chosen depending on the specific needs and on the resources available for evaluation by the host, the designer, or external evaluators. These elements might be used to generate more data on the value of laboratories and help answer open research questions in the future:

1. During the initiation of the event and more intensely during the co-design, establish evaluation as an integral part of activities and resource allocation.
2. During the co-design phase, create clarity about the objectives and intentions of the local champion, the designers, and the participants. Cross-check with the four areas of acquisition, association, awareness, and application. Be transparent about the main intentions in the invitation sent to participants.
3. Before the event, find out more about the interests of the participants and their current competence in using the future through a survey. This can help to establish a baseline for the evaluation. In addition, the results could inform the design and the style of facilitation of the event.
4. At the start of the event, ask participants about why they came and what their hopes are for the event. The overlap or not with the intentions formulated in the invitation may provide valuable insights for introductory remarks and the overall facilitation.
5. During the event, observe participants' behavior, engagement, and reactions. Discuss those in the hosting team and adjust the design if needed. Take notes.
6. At the end of every event and potentially also during longer events, ask participants to reflect on their experience and to share their reflections in written or oral form. If in oral form, one team member should take notes for later analysis.
7. A survey could be administered just before the closing to get a high response rate. This survey could ask questions about the overall design of the laboratory, whether the intentions were met, and about the quality of facilitation.
8. After the event retrospective meetings could be held with the design team and with the local champion, maybe using a structure that relates to the intentions agreed upon during the co-design phase. A report about the insights could be written for future reference.
9. After the event, interview a sample of participants or key stakeholders about their experience, again relating to the original intentions.

-
10. Track and collect feedback that comes in later on a scattered basis. Are participants asking for more information about the method? Did they recommend the method to others? Are they doing something with the results?
 11. One or two years later, interview participants again about what they recall from the event and whether they applied any of the insights or the method.

In line with Cabaj (2014), space should be given in each feedback option to articulate results that may not have been intended. In general, the need of measurement for evaluation should not limit the opportunities for emergence during the laboratory itself.

Evaluating interventions in complex systems takes time and effort. And it requires some skills that are similar to the skills of facilitators of Futures Literacy Laboratories. Gamble, McKegg and Cabaj (2021) summarize those skills from a developmental evaluation perspective. They include the ability to raise evaluative questions, which are valuable and important for those using the evaluation. Evaluators also need strong pattern recognition skills, which is a general requirement for anyone operating in complex systems. Thirdly, they need the capacity to frame the situation by identifying and naming what is happening. Fourth, since evaluation is part of the intervention and requires a lot of interaction with others, the ability to work in teams is crucial for every evaluator. Fifth, evaluators must be comfortable and familiar with a range of methods and assemble them in the appropriate constellation for the intervention at hand. Also, the ability to navigate conflict, power, and uncertainty is important.

If the evaluator is a member of the core team, his or her neutrality might be an issue to keep in mind, as well as a possibly limited knowledge of evaluation methods (Gardner, 2017). On the other hand, if a neutral and experienced evaluator is brought in from outside, he or she may lack familiarity with the way of working in a Futures Literacy Laboratory. A team approach might combine strengths.

5.3 Results and use of evaluations

Specific evaluation results did not play a major role during the interviews. Those that were mentioned did map into the intentions outlined earlier and into the general design principles of laboratories. For example: “Participants express issues and see things they never expressed before, including unpleasant ones.” Also, participant reactions include expressions such as eye-opening, mind-blowing, surprise, sense of wonder, and co-creation, which relate to the awareness of novelty. One inquiry into what participants remembered two years after an intervention revealed relationships, the positive space, and the sculpture co-created in the reframing phase.

The final question about the use of the insights from the evaluation led to the expected results mentioned in the introduction: Designers and facilitators use them to adjust the laboratory in real time. Also, insights are used to report to the local champions so that those can better justify the investment of resources. The use of the insights for the design of future laboratories was not mentioned explicitly, possibly because it appeared to be obvious.

Regarding unintended consequences, one interviewee raised ethical issues related to the impact of Futures Literacy Laboratories. If participants saw more opportunities than they did before, moved outside their old boxes, and self-reflected on their values, what do they do with this after the laboratory? What impact might this have on their private and professional lives?

6. Discussion

This paper was motivated by the observation of limited evaluation elements in futures work in general and on Futures Literacy Laboratories in particular. This observation contrasts with the many potential benefits of evaluation in terms of increased credibility among users and a better understanding of what works in what context. A lack of funding for internal or external evaluation capacity appears to be one reason for this contrast. Another reason might be a general skepticism toward evaluation in complex emergent systems. This paper suggests that this skepticism is unjustified. The main insights from this research are:

1. It is possible to evaluate interventions in complex emergent systems, in particular collective intelligence knowledge creation processes. There is a large body of valuable theoretical research and practical experience.
2. A surprisingly large amount of evaluation is already done on Futures Literacy Laboratories using a multitude of tools roughly in line with the insights from evaluation in other disciplines.
3. From the different data points made visible in this research project, patterns of reasons for evaluation, specific targets, practices, and uses of evaluation have become visible.

Evaluation and the use of Futures Literacy Laboratories have the potential for going hand in hand: Wider use of the method can create more opportunities and ideally more resources for evaluation. And supportive evaluation results could lead to a wider and more tailor-made use of the method by convincing potential users of the value of this intervention.

Several open issues emerged during the project and could be the subject of future research, especially as the database on evaluation expands. For example, different intentions of local champions, participants, and designers could have effects on the co-design of laboratories, the experience of participants, and the likelihood of local champions using the method again later. Future research could also investigate whether designers' knowledge about participants is indeed as limited as indicated by the ten interviews in this project. And it could try to find out whether designs and outcomes would be different depending on the depth of prior knowledge about participants.

Regarding how the evaluation is done in practice, future research could collect and develop appropriate questions for surveys that can reveal emotions. And there could be an investigation of how surveys might be used as an integral part of the intervention. If surveys are used more widely, it may also become possible to set up a collection of survey questions before, during, and after laboratories. This could include the context of their use, the intention behind it, and the value they created.

Similarly, on the use of participants' self-reflection: If this kind of activity is used more widely for evaluation, future research could focus on possible framings for reflective work to limit the effect of social acceptance. There could be learning potential from neighboring disciplines. Regarding the retrospective meetings with the team and with the local champion, there could be research on useful structures and whether the meetings can be informed by retrospective practices related to Scrum or Agile.

References

Aaltonen, Mika. 2009. "Evaluation and Organization of Futures Research Methodology – Version 3.0." In *Futures Research Methodology – V3.0*, edited by Jerome C. Glenn, and Theodore J. Gordon. The Millennium Project.

<https://www.millennium-project.org/publications-2/futures-research-methodology-version-3-0/>

Bressers, Nanny, and Lasse Gerrits. 2013. "A Complexity-Informed Approach to Evaluating National Knowledge and Innovation Programmes." *Systems Research and Behavioral Science*, 32(1), 50-63.

<https://onlinelibrary.wiley.com/doi/10.1002/sres.2205>

Burns, Anne. 2015. "Action Research." In *Handbook of Research in Second Language Teaching and Learning* edited by Eli Hinkel, 187-204. Routledge.

Burns, Danny. 2014. "Assessing Impact in Dynamic and Complex Environments: Systemic Action Research and Participatory Systemic Inquiry." *CDI Practice Paper 8*.

<https://opendocs.ids.ac.uk/opendocs/handle/20.500.12413/4387>

Cabaj, Mark. 2014. "Evaluating Collective Impact: Five Simple Rules." *The Philanthropist*, 26(1), 109-121.

<https://thephilanthropist.ca/original-pdfs/Philanthropist-26-1-17.pdf>

Cevolini, Alberto. 2016. "The Strongness of Weak Signals: Self-Reference and Paradox in Anticipatory Systems." *European Journal of Futures Research*, 4(4), <https://doi.org/10.1007/s40309-016-0085-1>

Chen, Kuo-Hua, and Li-Ping Hsu. 2020. "Visioning the Future: Evaluating Learning Outcomes and Impacts of Futures-Oriented Education." *Journal of Futures Studies*, 24(4), 103-116.

<http://jfsdigital.org/wp-content/uploads/2020/06/11-Chen-Visioning-the-Future-ED-8-2.pdf>

Cilliers, Paul. 1998. *Complexity and Postmodernism – Understanding Complex Systems*. Routledge.

Dator, Jim. 2019. "Editorial: On Evaluation and Foresight." *World Futures Review*, 11(4), 292-293. <https://doi.org/10.1177%2F1946756719867226>

Eoyang, Glenda H., and Thomas H. Berkas. 1999. "Evaluating Performance in a Complex Adaptive System." In *Managing Complexity in Organizations*, edited by Michael R. Lissack, and Hugh P. Gunz. Greenwood Publishing Group.

Canadian Evaluation Society. 2015. What is Evaluation?

https://evaluationcanada.ca/sites/default/files/ces_def_of_evaluation_201510.pdf

Gamble, Jamie, Kate McKegg, and Mark Cabaj. 2021. *A Developmental Evaluation Companion*. The McConnell Foundation.

<https://mcconnellfoundation.ca/wp-content/uploads/2021/06/A-Developmental-Evaluation-Companion-McConnell-Foundation-2.pdf>

Gardner, Annette L. 2017. “Designing Advocacy and Policy Change Evaluations”. In *Advocacy and Policy Change Evaluation: Theory and Practice* edited by Annette L. Gardner, and Claire D. Brindis, 75-113. Stanford University Press.

Gardner, Annette L., and Peter Bishop. 2019. „Expanding Foresight Evaluation Capacity.” *World Futures Review*, 11(4), 287–291.

<https://doi.org/10.1177%2F1946756719866271>

Inayatullah, Sohail, ed. 2004. *The Causal Layered Analysis (CLA) Reader*. Tamkang University Press.

Kauffman, Stuart. 1995. *At Home in the Universe: The Search for the Laws of Self-Organization and Complexity*. Oxford University Press.

Kazemier, Elles, Loes Damhof, Jitske Gulmans, and Petra Cremers. 2021. “Mastering Futures Literacy in Higher Education: An Evaluation of Learning Outcomes and Instructional Design of a Faculty Development Program.” *Futures* 132. <https://doi.org/10.1016/j.futures.2021.102814>

Maurer, Martina. 2021. *Der Mehrwert von Labs als Ansatz in Führungs- und Systementwicklung*. Springer Gabler.

Miller, Riel, ed. 2018. *Transforming the future – Anticipation in the 21st Century*. UNESCO and Routledge. <https://doi.org/10.4324/9781351048002>

Patton, Michael Quinn. 1978. *Utilization-Focused Evaluation*. Sage.

Patton, Michael Quinn. 2010. *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use*. The Guilford Press.

Patton, Michael Quinn. 2019. “Expanding Futuring Foresight through Evaluative Thinking.” *World Futures Review*, 11(4), 296–307.

<https://doi.org/10.1177%2F1946756719862116>

Poli, Roberto, ed. 2007. *Handbook of Anticipation – Theoretical and Applied Aspects of the Use of Future in Decision Making*. Springer.

<https://doi.org/10.1007/978-3-319-31737-3>

Rosen, Robert. 2012. *Anticipatory Systems: Philosophical, Mathematical, and Methodological Foundations*. 2nd edition. Pergamon Press.



This publication is available in open access under the license “Attribution-ShareAlike 4.0 International”