

CASE STUDY

REPORT

Monitoring and
evaluation practices:
UIA lessons learnt

PROJECT

U-RLP Utrecht Refugee
Launch Pad (Closed)

📍 Utrecht, The
Netherlands

TOPIC

Integration of migrants
and refugees

EDIT 04 MARCH 2021

Utrecht U-RLP

About U-RLP

See on UIA
website



© Henni Bunnik

The 'Utrecht Refugee Launchpad' (U-RLP), which came to be known as 'Plan Einstein', was implemented by the city of Utrecht and its partners between November 2016 and October 2019. As an innovation, the project aimed to address three major problems identified in solutions related to the reception of asylum seekers. According to the [final evaluation report](#), these included:

- Limited opportunities for building community cohesion;
- Arrested labour market activation;
- Poor wellbeing of asylum seekers.

Plan Einstein expected to generate better relations in the neighbourhood, activation for newcomers from day one and continuity for refugees given status, ultimately creating better wellbeing for asylum seekers in the city.

Source: [Final evaluation report](#)

The project targeted asylum seekers from the Overvecht asylum seeker centre and members of the neighbouring community. It consisted of three main types of activities: co-housing, co-learning and individual support for asylum seekers. As part of co-housing, young tenants moved into subsidised rooms adjacent to the asylum seekers' centre. A common incubation space was created in the building for use by tenants, asylum seekers and other community members. Through co-learning, asylum seekers and community members benefited from courses in English and entrepreneurship, considered as 'future-proof' skills. Individual support for asylum seekers aimed to encourage their participation, but also reduce stress levels and the consequences of experienced trauma. Additional opportunities to both learn and integrate were also provided. More information about the activities can be found [here](#) and in the [final evaluation report](#).

Evaluation governance

In the overall project design, the evaluation constituted a separate workstream with dedicated partners. It was planned as an independent exercise to be conducted over three years by researchers representing two UK-based academic centres – University College London (UCL) and the University of Oxford.

The study was led by an expert from UCL trained in sociology and anthropology, supported by a small team. In addition to a trained research team, the evaluation benefited from the expertise of an academic Advisory Board, chaired by a representative from the University of Oxford. The board gathered distinguished scholars specialised in issues relevant to the project, such as the reception of asylum seekers, but also to evaluation itself. It met three times – to discuss the methodology, interim findings and draft final report.

That group which, in a sense, had oversight of the research, but no control of it, just advised on it [...] It was the one forum which looked at the research exercise from an academic perspective [...] Unconflicted by any other issues.

Source: *U-RLP project hearing*

Both the evaluators and the project management found the Advisory Board to offer important support. With its multidisciplinary in-depth academic expertise and independence, the Advisory Board was able to provide subject-matter expertise, but also challenge assumptions (underlying the theory of change, for example) or expose possible biases.

To have that external body was very helpful, because it really stopped that pro-innovation bias [...] We had to work quite carefully to make sure that – you know – the message was constructive and helpful, but actually having the Advisory Board there was good really to help us push back when we needed to.

Source: *U-RLP project hearing*

It was also in a position to provide guidance to the researchers in a dynamic evaluation context, which required flexibility in the methodology and added nuance to the findings.

While the project beneficiaries were involved in the evaluation through a large number of individual interviews, ideally, the lead evaluator would have liked to involve representatives of the target group in data collection as community researchers. This was not possible in the project's short timeframe, however.

The project created cooperation mechanisms which enabled implementation of the evaluation, on the one hand, and feeding of the evaluation results into the partnership on the other. The evaluation included extensive cooperation with all project partners at various stages. They were included in the development of the approach and the project's **theory of change**. They supported the data collection process and were informed on the evaluation's progress. The evaluators, while not being involved in the main activities, closely followed the project's work and participated in its various meetings. Specific meetings were dedicated to discussions related to the evaluation itself. One meeting was specifically devoted to extensive explanation and clarifications of the applied approach for project partners. Separate meetings were also organised to present and discuss evaluation results, both interim and final.

The experiences of U-RLP's evaluation show how important it is to set aside appropriate resources for evaluation in order for it to produce expected learning. As the project evaluators noted, the aspiration to measure impact is an ambitious one, so it should be allocated an appropriate budget and timeframe.

Ultimately, the value of the project is in what you learn from it. It was an awful lot of value for the people who were affected, but that passes and then the ultimate value is in what you learn, so

evaluation shouldn't be an afterthought in terms of resources, but more central to that.

Source: *U-RLP project hearing*

If appropriate resources are not available, then expectations of what can realistically be delivered should be managed.

U-RLP's experience underscores the need for sufficient time to be earmarked in two dimensions in particular – working time for researchers and time for the evaluation itself, including for its preparation. The project's evaluation can be seen as an example of best practice for various reasons, and both its breadth and depth are impressive. This was possible thanks to the in-depth expertise, as well as strong personal engagement and commitment, of the otherwise rather under-resourced evaluation team, supported by project management.

We found this project absolutely mushroomed. And I had the equivalent of research time, including myself, of one full time employee working on this. To do and to deliver everything that we did only happened because we all put in way, way more time.

Source: *U-RLP project hearing*

One of the reasons why resources were insufficient related to the fact that not all obstacles and difficulties were anticipated from the start. As the lead evaluator observed, had there been more clarity in this respect, she would have budgeted for more research time. It is, thus, an argument for devoting a sufficient amount of time for risk analysis at the evaluation design stage.

Apart from working time, U-RLP experience also shows that the project would have benefited from a preparatory phase. This would have allowed for a more comfortable development of a theory of change (see also below a discussion on how time-consuming this process was) and research tools prior to the implementation of activities themselves. The lead evaluator estimated such a phase to require roughly a couple of months, possibly as many as six. Importantly, such a preparatory phase within projects was introduced by UIA in subsequent funding editions.

Theory of change

"A 'theory of change' explains how activities are understood to produce a series of results that contribute to achieving the final intended impacts. " A 'theory of change' should, therefore, present the link or path between what one is doing and what one is trying to achieve.

Source: Rogers, P., [Theory of Change: Methodological Briefs - Impact Evaluation No. 2](#) Methodological Briefs no. 2, 2014.

Evaluation process

General approach

U-RLP's evaluation is an example of a comprehensive and coherent approach which responded to the project's reality. The initial expectation was that the evaluation would follow an experimental or quasi-experimental approach. This would have helped determine causal relations between participation in the project and specific results or impacts. However, the project's nature and complexity, as well as a number of resulting considerations, made implementation of such an approach impossible, even though it had seemed feasible at the time of designing the evaluation.

When you're designing these things they are on paper and [...] it's very difficult to see how it actually works. [...] then the funding came through and we started talking through what was possible. It just seemed non-experimental was probably the more appropriate design [...] because of [...] the nature of the population, for a start.

Source: *U-RLP project hearing*

Experimental and quasi-experimental design

Experimental and quasi-experimental designs aim to test causal hypotheses, e.g. that a given change resulted from the project. Experiments differ from quasi-experimental designs in the way subjects are assigned to treatment. In true experiments, subjects are assigned to treatment randomly, e.g. by a roll of a dice or lottery, to eliminate selection bias. After the process of random selection, selected groups are then subjected to identical environmental conditions, while being exposed to different treatments (in treatment groups) or lack of treatment (in the control group). A quasi-experimental design lacks random assignment. Instead, assignment is done by means of self-selection (i.e. participants chose treatment themselves), administrator selection (e.g. by policymakers, teachers, etc.), or both.

To learn more about this approach, you can consult e.g.:

- White, H., Sabarwal, S., [Quasi-Experimental Design and Methods](#), UNICEF, 2014

General approach

The [final evaluation report](#) lists a number of challenges which prevented application of the experimental approach, namely that:

- Participants experienced the project in different ways, to varying degrees and for different periods of time, which made comparisons difficult;
- Random assignment to experimental and control groups, i.e. provision of support to some asylum seekers and denial to others based on random selection, was highly problematic ethically in a project which targeted a vulnerable population;
- Asylum seekers themselves could choose to participate or not to participate in the project (or in parts of it);
- Due to varying conditions in asylum seeker centres in the country, it was hard to find any centre which could function as a comparable control group for the experiment.

Application of an experimental design was also hindered by the decision of the Dutch asylum authority (COA) to deny the evaluation team access to asylum seekers residing in the centre unless they participated in the U-RLP project. This meant that the evaluation team did not have access to those asylum seekers from the Overvecht centre who did not decide to be involved with the project. Consequently, the sample of accessed asylum seekers was biased towards those receiving support and the opportunities for comparison with those who did not receive it were limited. More information on the considerations which underly U-RLP's decision not to pursue an experimental design can be found in the relevant part of the [final evaluation report](#). In this respect, the U-RLP intervention shares similarities with the evaluation conducted within [Rotterdam's BRIDGE project](#).

Eventually, the evaluation followed a theory-based approach, involving the use of a theory of change for both designing and implementing the evaluation and determining the project's contribution to expected outcomes. As the final evaluation notes, this approach "integrated rather than eliminated these important contexts", which posed a challenge to the experimental approach "by seeking understanding of the interplay of project and effects". In other words, instead of causal relations, this approach allowed evaluators to focus on the questions of how and why things had happened.

Rather than seek to prove causal attribution, the goal of the evaluation was to explain, with sensitivity to contexts, how the existence and operation of Plan Einstein likely contributed to observed outcomes, or made 'a difference'.

Source: *U-RLP Final evaluation report*

Theory-based approach

Theory-based evaluation has at its core two vital components. Conceptually, theory-based evaluations articulate a policy, programme or project theory, i.e. how activities are supposed to lead to results and impact, given specific assumptions and risks. Empirically, they seek to test this theory, to investigate whether, why or how interventions cause intended or observed results. Testing the theories can be done on the basis of existing or new data, both quantitative (experimental and non-experimental) and qualitative.

Source: European Commission, [Evalued Sourcebook – Method and techniques](#), 2013. For more information, visit e.g. [Better Evaluation website](#) or the website of the [Treasury Board of Canada Secretariat](#).

The U-RLP project did not have a theory of change from the start, so theory development was the first stage in its evaluation. This helped to define what the project was trying to deliver and to bring partners towards a common understanding of these aspirations. Theory of change proved to be a flexible enough methodology to enable collaborative work and allow space for different points of view. As this suggests, since it can help to define

objectives and build agreement between diverse stakeholders, the approach can be particularly useful in projects which involve many partners. Similarly, it can fit those interventions which were designed by one group of people yet implemented by another. Both circumstances were present in the U-RLP project. Importantly, with academic partners leading the evaluation, the theory developed with U-RLP project partners was also rooted in the latest science related to asylum seeker reception, and their integration with local communities and the labour market.

Theory of change was identified as an appropriate thing to try and clarify some of the thoughts that were going around about what this project was going to deliver. In the application [...] there is a lot of aspirational language, but actually pinning down what the results are is quite difficult, particularly when you have so many different partners who were coming at things from their own vantage points.

Source: *U-RLP project hearing*

As U-RLP's experience shows, theory development also posed some challenges. In particular, it proved very time-consuming as a process in a multiple partner setup. The innovative and pilot character of the project, which meant less clarity about the aims and possible outcomes at its outset, intensified this challenge. For this reason, the project evaluators underscored the importance of allowing an appropriate amount of time for this stage of the evaluation process, and for evaluation preparation more generally. This is particularly true if the theory of change is to be consistently used for monitoring the project, as well as designing and implementing the whole evaluation.

If you want to do an evaluation of whether the aims have been achieved, you've got to have a clear understanding of what the aims are. And starting out with, I think, 57 aims – that's problematic. [...] And my recollection is just how much time you've had to spend working through with everyone, say what is it actually you really are trying to achieve here, let's narrow down the 57 a bit.

Source: *U-RLP project hearing*

Another challenge in implementing the approach related to explaining the meaning of theory development to project partners who did not necessarily have a research or evaluation background.

I was working directly with individual project partners to develop a theory of change. We then had a workshop on it. And then I was sending things out for feedback. But, if I'm honest, I don't think everyone understood what I was trying to do.

Source: *U-RLP project hearing*

This shows that in complex, multi-partner interventions, evaluators need to also be good communicators who are able to explain the meaning of various evaluation processes and, project-wise, bring all the partners to a common denominator. This is easier when evaluators have enough platform in the project to interact with other partners, as was the case in the U-RLP project.

While theory-based evaluations sometimes face criticism for their actual limited use of the developed theory, Plan Einstein employed its theory extensively throughout the project and evaluation. Thus, the theory served as basis for project monitoring. It was also used to develop evaluation questions and determine data collection methods and sources, as well as to develop coding for data analysis.

Importantly, however, the U-RLP lead evaluator felt that there was potential for pushing the approach even further. For interventions which consist of various workstreams, each devoted to different activities which beneficiaries can choose from, she felt it may be useful to develop nested theories of change. This means creating lower-level theories of change for specific project components. Such a solution could help evaluators to better distil the outcomes of particular components and show the workings of the project at a more granular level.

The U-RLP evaluation was "inspired by the idea of learning through the process, providing not only accountability, but drawing out lessons that can be taken from the experience of innovation." ([Final evaluation report](#)). Importantly, however, while the evaluators were inspired by the idea of learning, they felt constrained by the need to ensure project accountability.

Because of the fact that the Commission wanted accountability as well as learning, we had to sort of straddle this role. And actually, they're kind of, both conceptually but also practically, quite different approaches that you would take.

Source: *U-RLP project hearing*

Espousing those two very different goals created some tension in the process. The need to preserve a component on accountability was the reason why the evaluators opted for a theory-based approach instead of a developmental evaluation approach, which they considered better suited for innovative projects such as U-RLP. The theory-based approach, even if collaborative, foresees determination of project outcomes at the beginning and, in this way, is not perfectly compatible with

projects of a pilot nature where these are still uncertain and being determined more dynamically.

If you're trying to retrofit a kind of more traditional evaluation approach onto this, even theory of change – which is still very collaborative – it is still trying to impose a linear logical progress to a set of activities which are untested, unknown and actually do need the time to develop. And as evaluators, our role should be in supporting that.

Source: U-RLP project hearing

Approach to data collection

To best capture evidence of early outcomes, the evaluation team proposed an elaborate mix-methods approach to the study. It used a combination of qualitative and quantitative methods for data collection and analysis, as well as a wide array of data sources summarised in the table below.

Quantitative research	Qualitative research
Routine monitoring of activities and participation	163 interviews with 127 interviewees (36 repeated)
Analysis of NOA intake assessments completed by asylum seeker participants (N=150)	Documentary analysis of minutes of meetings and consideration of supplementary material associated with the project (advertisements, flyers, reports, photographs)
Class and activity evaluations: Questionnaires co-designed with partners that were administered at the end of the class series (111 responses) and online surveys for participants in the business programme (95 usable responses)	Media analysis of Dutch and English newspaper sources between beginning of January 2016 and March 2019, retrieved through Nexis Uni and TV items collected through the database of the Netherlands Institute for Sound and Vision
Analysis of wider available quantitative data, from stakeholder organisations and the local government	

Data collection methods

Source: Final evaluation report

One of the challenges from the perspective of determining the project impact was securing an appropriate baseline, i.e. an overview of the situation prior to the intervention's initiation. Even though the evaluation was started immediately, with no overall preparatory phase which would allow for development of research tools, the first evaluation measurements were delayed in relation to other activities.

We've already said about the fact that the baseline was compromised, so that was obviously a problem that we were constantly playing catch up from the word go, with everything actually.

Source: U-RLP project hearing

The measurement of attitudes, aspirations and hopes for the project among young Dutch tenants was conducted after they had already moved into their apartments. It would have also been more valuable if the neighbourhood survey had been carried out at the time when the hostilities towards refugees were high. Consequently, the **baseline measurements** offered an already altered picture of the intervention, compared to its actual starting point.

In light of the sensitive nature of the project, the implementation of the neighbourhood survey had to be approached with caution, specifically in relation to its language and distribution. The idea of sending the surveys out via post was eventually rejected, since it could have opened the results to hijacking and manipulation by groups hostile towards the asylum seekers' centre in the neighbourhood. Instead, a face-to-face option was selected and implemented by trained researchers.

With the neighbourhood survey [...] we were asking what they thought about asylum seekers, which had been an inflammatory issue [...]. So, it was incredibly important to get that right in the approach, so actually thinking through the questions, about using this in a way that wouldn't provoke, but also that wasn't open to abuse.

Source: U-RLP project hearing

Mindful of the risk of overburdening refugees with research activities, during the research, the evaluation team decided to use the available NOA intake assessments, i.e. pre-existing quantitative baseline assessments completed by asylum seekers when joining the project. The initial assumption was that these would be later replicated for evaluation purposes to offer a comparison with the baseline assessments. However, repeating NOA assessment proved impossible during the evaluation. Asylum seekers perceived the surveys as a test and were likely to present a desired picture, which raised both ethical and validity questions for the evaluation. Therefore, a decision was made to pursue a qualitative approach to collecting data with asylum seekers through individual interviews.

The large component of individual interviews with foreigners activated considerations related to translation and interpretation, including the specificity of working with interpreters in sensitive, qualitative research. The evaluation team was aware of the power relations between interviewees and interpreters involved in such processes. For similar future studies, they would have therefore preferred involving community researchers. Not only would that reduce the power dynamic, but it would also increase the evaluation's participatory character and produce additional opportunities from project beneficiaries employed as such researchers.

It's always difficult when you're doing research with refugees in multiple languages. On occasion, we were using translators. I'm aware [...] of the power dynamics of that [...] But, because of the nature, the convenience, we had to do these things [...] I would prefer to [...] use community researchers. I don't know how possible it would be to do. The brevity of this project in the end meant that we couldn't really do that properly.

Source: U-RLP project hearing

It is clear that despite ethical, political, temporal, logistical or language-related limitations, the evaluation design allowed the team to gather a wealth of data from different sources and give strong evidence backing for evaluation claims.

Horizontal issues

In projects such as U-RLP, ethical and political considerations play a significant role in the evaluation. At the project outset, the vulnerability of the targeted population added a layer of complexity to the university ethical review. As recalled by the lead evaluator, the review was incredibly detailed and, consequently, time-consuming. In this respect, yet again, the project would have benefited from a development stage to account for that period.

Ethical considerations, among others about not burdening the asylum seekers with research activities, also influenced methodological changes, such as resignation from replicating the quantitative NOA assessments in favour of conducting qualitative individual interviews. They also played a part in the development and conduct of the neighbourhood survey. Considering the highly controversial subject, the survey questions had to be drafted with particular care so as not to stir controversy in the neighbourhood.

Political factors also had an influence on the project methodology. With respect to the survey, they influenced its distribution towards a face-to-face mode. The Dutch asylum authority's refusal to provide the evaluation team with access to the asylum seekers' centre prevented the researchers from talking to asylum seekers who did not benefit from the support offered as part of the U-RLP project. This meant that comparing the situation of U-RLP beneficiaries to Overvecht asylum seekers who did not receive support was not possible.

Lessons learnt

The U-RLP offers a rich source of insight for those embarking on an evaluation of a similar complex intervention related to serious social challenges. Many of those have been noted in previous parts, and below we additionally summarise our selection:

There should be time assigned for development of the project and its evaluation. The provision of such time would have allowed the evaluation team to recreate the theory of change, develop research tools and carry out appropriate baseline measurements prior to the initiation of actual project activities. Without such a phase, the baseline was compromised and some opportunities for capturing the project's contribution to outcomes

potentially lost, while the evaluation team were playing catch up from the word go.

Theory-based evaluation and theory of change can be useful, yet possibly not ideal for innovative interventions. The selected approach worked in practice and offered a good opportunity for collaboration and inclusion of various points of view. The theory of change was used consistently and productively throughout the whole evaluation. At the same time, given the project's complexity and with more time available, the evaluators would have liked to develop nested theories of change for particular project components. In view of the pilot character of the project, they would have generally preferred a more flexible approach, such as developmental evaluation, which does not impose a linear logic onto a dynamic innovative intervention. This would have helped to concentrate resources on generating insight for the project as to how and why things worked the way they did.

There is a need for significant resources. The ambitious goal of capturing early outcomes and impact should receive appropriate resources, both in terms of funding and time. In the space of the U-RLP project, budgetary shortcomings were compensated for with individual engagement of researchers. However, this cannot be a standard expectation for professionals who engage in evaluation. Time-wise, the funders interested in measuring impact should allow the evaluation to reach beyond the project implementation time.

If you want to capture outcomes – what were the long-term implications for the community and for the individuals involved – you need to carry on after the end of the initiative, or the project itself needs to carry on for longer.

Source: *U-RLP project hearing*

Academic expertise is an asset. The U-RLP project could rely on the extensive expertise of its academic partner both at the level of the core evaluation team and the evaluation Advisory Board. U-RLP evaluation experience shows the importance of subject-matter expertise and research skills. These offer backing and add nuance to various elements of the research, be that the theory of change, methodological choices made throughout the study or its specific findings. This rich research experience allowed the team to develop a complex evaluation design, involving a dynamic and creative approach to data collection, to compensate for the lack of a counterfactual and to respond to various ethical and political challenges along the way.

The evaluator's role can go beyond passing judgement on the project and their position may require balancing between different objectives and expectations. The U-RLP experience shows that in complex, multi-partner projects, evaluators need to be good communicators who are able to explain the meaning of various evaluation processes and, project-wise, bring all partners to a common denominator. Some tension was also revealed in the evaluator's role between the need for accountability and the need for learning through the evaluation. This tension created a level of discomfort, and the need for accountability inevitably diverted resources from learning opportunities. With innovative interventions such as U-RLP, evaluations would have benefited from concentrated focus on generating insight and understanding about projects' inner workings and achievements.

The final point I would make is about this role of the evaluator. I think it needs to be made very clear that it is not an accountability role [...] It's to pull out the lessons learnt from it [...] The researcher's job is to tease out a more nuanced understanding of what worked and what didn't work.

Source: *U-RLP project hearing*



[See on UIA website](#)

