

## 3 Methodology

This chapter summarises the GOAL evaluation methodology, discussing the overarching evaluation design, quantitative and qualitative methods of data collection and analysis, and methodological challenges across the six-country project as a whole.

The GOAL **evaluation had three aims**. The first was **developmental**: to support programme development across the six countries by providing evidence during the life of the pilot on **programme processes**. The second was **summative**: to assess, as rigorously as possible, **the impacts of GOAL** on service users and other programme stakeholders. The third aim focused on **knowledge cumulation**: to provide evidence on programme processes and outcomes in order to support future policy and programme development in the field of adult education guidance. As part of this aim, we sought to help advance programme theory in this under-researched area. In doing so, the evaluation may potentially offer evidence and analysis that can be used by policy makers in similarly complex fields, to help them better understand challenges and success factors in establishing “joined-up” programmes in complex policy fields.

### 3.1 Evaluation design and methods

The methodological approach for this evaluation was shaped by the complexities of the project design, namely that:

- GOAL was multi-site (two sites in each of five countries, and four sites in the Netherlands) and multi-organisational.
- GOAL had multiple objectives.
- GOAL was predicated on cross-organisational collaboration.
- Each partner country had its own unique context and target groups (and target numbers to achieve).
- Programme resources were finite, and were by necessity and logic primarily focused on the interventions rather than the evaluation.

For these reasons, it was neither feasible nor advisable to conduct an experimental or quasi-experimental evaluation involving comparison groups. Instead the evaluation has positioned itself within the broad ‘**Theory of Change**’ (Weitzman et al., 2002) approach. Evaluations adopting this approach (Rogers and Weiss, 2007; Weiss, 1997) typically seek to address two levels of theory: 1) Implementation theory and 2) Programme theory.

**Implementation theory** focuses on how programmes are implemented, e.g. the intervention strategies that underpin programme activities. **Programme theory** focuses on programme *mechanisms*, by which we refer not to programme *activities* but to the *changes* within the participants that those activities facilitate. These changes, in turn, may lead to the desired

programme outcomes. For example, in a counselling programme such as GOAL, counselling is not a mechanism, it is a **programme activity**. Programme activities will ideally **trigger mechanisms (i.e. responses) within programme participants** – such mechanisms may include greater knowledge, increased confidence or motivation, and/or heightened ambition. These mechanisms, in turn, may then contribute to client actions and outcomes, such as enrolling on a course. As these examples illustrate, mechanisms serve as potential *stepping stones* towards the achievement of desired programme outcomes. Theory of Change approaches to evaluation focus not just on outcomes (i.e. what happens) but on how these outcomes are achieved. As such, there is a strong focus on these stepping stones.

While drawing on Theory of Change approaches in general, the GOAL evaluation also draws on a specific type of Theory of Change evaluation: **Realist Evaluation** (Pawson and Tilley, 1997). The Realist approach emphasises the central importance of the **interplay between programme contexts and mechanisms (stepping stones)**. A central tenet of Realist Evaluation is that programmes do not themselves produce outcomes in a direct causal fashion: programmes are not catapults with which clients are metaphorically launched into a better future. Unlike balls launched by catapults, clients have **agency**. Furthermore, they live their lives within **structural contexts**; these contexts produce **constraints and opportunities** within which agency may flourish (or not) to greater or less degrees. Causality (in terms of the intervention producing the desired effects) is thus contingent rather than deterministic: in the appropriate context and for the people, programmes (through their activities) *may* facilitate the triggering of mechanisms which *may* in turn lead to desired outcomes. Realist Evaluation, as with Theory of Change evaluation more generally, seeks to develop and test hypotheses about which interventions (or aspects of those interventions) work for whom in what contexts. As a corollary of this objective, Realist Evaluation rejects the assertion that to be considered successful, programmes must be context-independent, in terms of their ability to produce desired outcomes through the same intervention strategies for all target groups across all contexts. Whereas such context-independence and broad-scale generalisability may potentially be achieved with simpler interventions, it is unlikely to be feasible with complex interventions such as GOAL. A key objective of Realist Evaluation (and Theory of Change evaluation more generally) is thus to produce **theoretical generalisations which future programme developers and policymakers can draw on** when developing interventions in their own particular contexts and for particular target groups. This means measuring not only the degree to which a programme does or does not work, i.e. the degree to which it produces the desired outcomes, but also generating knowledge about **how programmes work, for whom, in what contexts, and why**. This requires in-depth understanding of intervention strategies and activities, and their relationship to programme contexts, mechanisms and outcomes.

In generating knowledge not just about whether programmes work but also how and why they do so, evaluators seek to go beyond merely providing a summative assessment of a specific programme. Summative evaluation is necessary but not sufficient. A broader goal is to contribute to the **cumulation of knowledge** in a field. Such cumulation, and the theory development it implies, is particularly essential in underdeveloped fields such as that investigated by GOAL: guidance and counselling for low-educated adults. This objective is important not just because of the **limited**

**amount of credible evidence in this nascent field**, but also because of the **inherently complex nature of interventions such as GOAL**. Evaluations which seek to understand and assess complex interventions must take account of a range of complicating factors within the programme (Rogers, 2008), including: 1) multi-agency governance and/or implementation; 2) simultaneous causal strands leading to desired outcomes; 3) alternative causal strands leading to desired outcomes; and 4) recursive causality. These four factors are discussed in the following paragraphs.

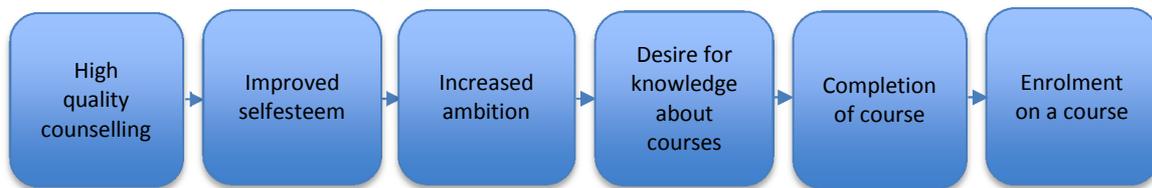
The importance of **multiple agencies** will be apparent throughout this report, particularly in discussions of partnerships and networks. As these agencies exist at programme and policy levels, the evaluation takes a multilevel approach: an important element of the evaluation is the description and assessment of the policy factors that play a role in influencing programme success. It is hoped that this dual focus on **programme-level and policy-level processes**, and their interaction, will provide useful evidence for a range of policymakers working in complex fields.

The notion of **simultaneous causal strands** refers to the presence of two or more causal strands that are required in order for desired outcomes to be achieved – e.g. for programme participants to enrol on a course, they may need to improve their motivation (causal strand 1), but viable courses also need to be made available to them (causal strand 2).

**Alternative causal strands** refers to the likelihood that one aspect of the programme may work for one client (in terms of producing a desired outcome), whereas another aspect may work for another client. For example, one GOAL client may take the “next step up” into education as a result of increasing their previously low self-confidence or self-belief. Another client may take the same step for a different reason, e.g. perhaps she was already motivated but simply lacked information about relevant courses.

Finally, the notion of **recursive causality** refers to the non-linearity of many causal pathways. A linear model of programme theory might, for example, show a client moving in a direct, linear fashion across the stages illustrated in Figure 3.1. overleaf.

**Figure 3.1.** Programme theory: linear model



As illustrated in this linear model, high quality counselling *leads to* improved self-esteem, *which leads to* increased ambition, *which leads to* desire for knowledge about further education courses, *which leads to* enrolment on a course, *which leads to* successful completion of the course.

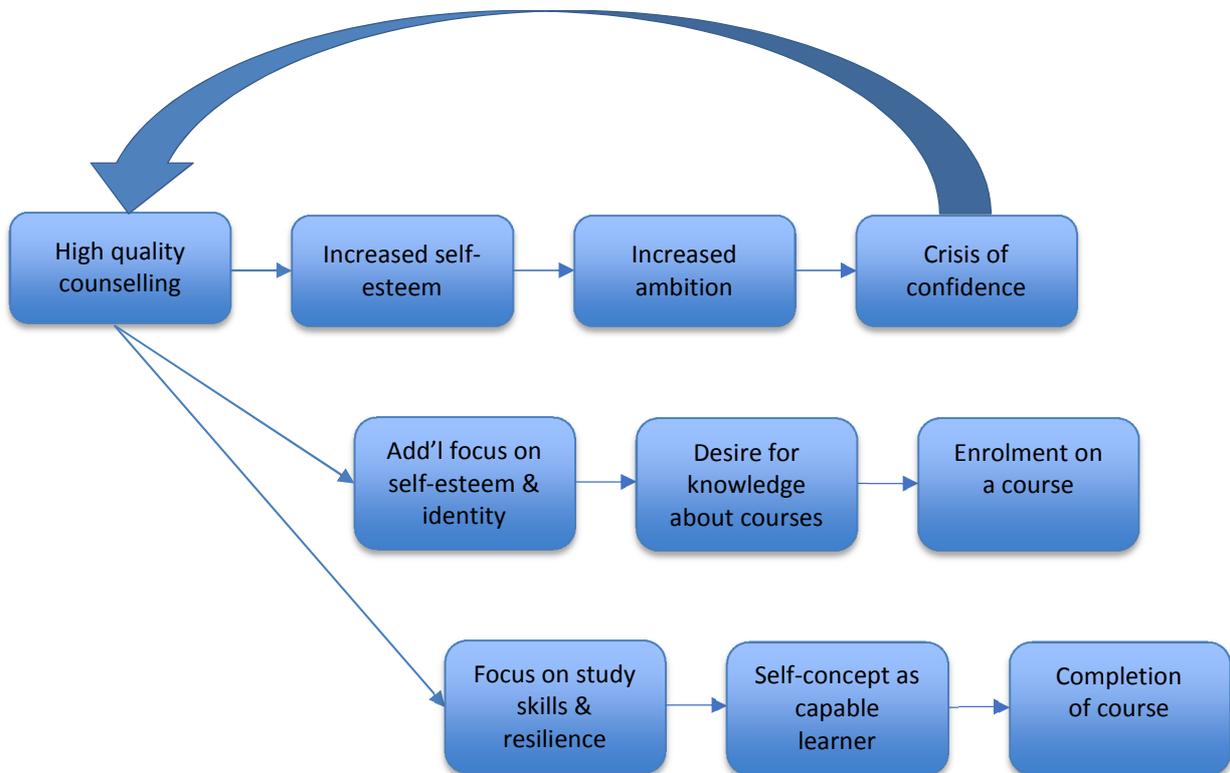
A more realistic (particularly for disadvantaged target groups), recursive model of causality might include all of these stages, but would take account of the tried and tested maxim that humans often need to take one step back in order to take two steps forward. Thus, a recursive model of causality (see Figure 3.2.) might be:

1. Improved self-esteem, *which leads to*
2. Increased ambition, *which leads to*
3. A crisis of confidence: the client had never seen herself as an ambitious person, and is uncomfortable or even threatened by this new identity. This *could lead to*
4. Additional focus on self-esteem and identity, *which leads to*
5. Desire for knowledge about further education courses, *which leads to*
6. Enrolment on a course, *which leads to*
7. Another crisis of confidence<sup>5</sup>, *which leads to*
8. Renewed focus on self-esteem and identity, plus a focus on study skills and resilience, *which lead to*
9. A new, expanded identity or self-concept as a capable learner, *which leads to*
10. Successful completion of the course.

A central objective of the GOAL evaluation is to develop and present a rich understanding of the range and types of causal pathways to be found in the programme, and the relationship of these pathways to specific national and local contexts.

<sup>5</sup> To reduce visual complexity this stage is not shown in the accompanying figure.

Figure 3.2. Programme theory: complex model



### Modifications to the Realist Evaluation approach

The Realist Evaluation approach has a straightforward (albeit analytically demanding) primary objective: to develop understanding of what works, for whom, and in what contexts. However, this straightforward objective is accompanied by rather non-straightforward academic jargon. Realist Evaluation terminology – particularly with regard to definitions regarding programme contexts, resources and mechanisms – can be difficult for non-specialists to grasp. In particular, there are conceptual challenges regarding what is meant by “context” and “mechanisms” (Dalkin et al., 2015). Rather than burden readers of this report with these challenges and the debates surrounding them, we have chosen to modify this terminology to make our analysis more easily understood by lay (non-academic specialist) readers. Instead of the tradition Realist Evaluation tripartite focus on contexts, mechanisms and outcomes, we have instead opted for a four-part approach, focusing on the following factors and their interaction:

- the **contexts** (e.g. policy environments) in which programmes are implemented
- the **resources and strategies** that programme staff draw on or utilise
- the intermediate or **stepping stone changes** (referred to as “mechanisms” in the Realist Evaluation literature) that lead to changes in action by clients or other programme; these stepping stones are changes in reasoning or beliefs that lead to programmes’ desired outcomes (e.g. an increase in client self-confidence is a stepping stone that may trigger an outcome such as enrolment on a course), and

- the interplay and influence of these contexts, resources, strategies and stepping stones on the achievement of **client outcomes** (e.g. enrolment on adult education courses) **and/or implementation objectives** (e.g. the successful development of partnerships and networks).

In summary, this evaluation has sought to achieve three **overarching objectives** (Berriet-Sollicec et al., 2014): 1) to measure the effects of GOAL, with regard to client outcomes; 2) to understand how, why, for whom and in what contexts outcomes are (or are not) achieved; and 3) to contribute to joint learning and knowledge cumulation – both (a) within the GOAL programme itself (e.g. by sharing process evaluation evidence with programme developers and other key stakeholders), and (b) in terms of the broader field of adult guidance and counselling (by providing credible and relevant programme theory and evidence that future programme developers and policymakers can draw upon in their own endeavours). In working towards these objectives, evaluation evidence has been gathered via:

- client monitoring data (to establish baseline, ongoing and exit data)
- client satisfaction and outcome data (user survey and qualitative interviews)
- programme and policy data (literature review; needs and strengths analysis)
- case studies of programme sites (qualitative interviews, document analysis, analysis of quantitative data)
- qualitative interviews with policy actors.

The evaluation includes: a) ongoing data collection (i.e. data collected throughout the life of the GOAL programme) and b) wave-specific data collection.

**Five research questions** underpinned the GOAL evaluation:

1. To what degree did programmes achieve their implementation aims across the five intervention strategies, and what factors at programme and policy level appeared to influence the achievement of implementation aims?
2. What service user outcomes were achieved, for what groups, and to what degree?
3. What programme-level factors were associated with the achievement of positive service user outcomes?
4. What policy-level factors were associated with the achievement of positive service user outcomes?
5. What was the Return on Expectations? That is, to what degree were programme expectations met?

### Return on Expectations (ROE)

Return on Expectation (ROE) analyses can support the development of programme theory by detailing programme stakeholders' (including clients, programme staff and policy funders) expectations of programmes and the rationales for those expectations, and then analysing the

degree to which those expectations have been met and the factors associated with that. ROE analysis can thus provide valuable programme and policy learning that contributes to the cumulation of knowledge in nascent fields, and helps in the development and refinement of future programme theories (Social Research and Demonstration Corporation, 2011).

### Local and central evaluation teams

As noted above, the GOAL programme was implemented in six countries, and an important consideration for the evaluation was to maximise learning about programme processes and outcomes at national level, as well as overall. That is, the research aim was not just to see if the intervention worked, but to understand and provide knowledge on the different ways that it worked (or struggled to work) in each of the six countries. A related evaluation objective was to provide stand-alone country-level evaluation reports as well as a synthetic cross-country report. For these reasons, a two-tier evaluation design was adopted: the overall GOAL evaluation was designed and run by the central UCL Institute of Education (IOE) evaluation team, but in each country there was a local evaluation team charged with collecting data and doing country-level analysis (guided by IOE).

### Interim (Wave 1) reporting

An interim national evaluation report for each GOAL country was published in November 2016. These reports, along with an interim cross-country report synthesising findings and key messages from all six countries, are available at <http://www.projectgoal.eu/index.php/publications> under the heading “Wave 1 Evaluation Reports”. A key aim of this interim reporting stage, which drew on data collected in Spring 2016 (see national reports for exact dates for each country), was to analyse and share early messages in order to facilitate service adaptation and improvement.

### Evaluation manual

To ensure the collection of robust data and the consistency of instrument administration across the six countries, IOE created an **evaluation manual** for use by local evaluators. Version 1 of the manual (November 2015) included protocols for two quantitative instruments used in ongoing data collection. Version 2 (March 2016) added guidelines for the administration of the Wave 1 Topic Guides and other instruments, as well as guidance for completing the interim national reports. Version 3 (February 2017) provided guidance on the development and use of all Wave 2 data collection instruments, and guidelines for completing the final national reports.

## 3.2 Quantitative sample and data collection

Due to **different implementation timeframes and challenges** across the six countries, client data collection began at different points: April 2015 (Flanders), January 2016 (Czech Republic), February 2016 (Iceland, Lithuania, and the Netherlands) and March 2016 (Slovenia). In all countries, data collection ended by April 2017.

Quantitative client data were collected throughout the life of the programme via: 1) a data monitoring instrument and 2) a client satisfaction survey. In addition, a follow-up survey was conducted with clients in Wave 2.

### Client monitoring data

The data monitoring instrument created by the IOE evaluation team gathered detailed information about the clients on the GOAL programmes, enabling evaluators to measure target numbers and track a range of programme processes and service user outcomes (see Appendix C, Data Monitoring Instrument, and Appendix D, GOAL data monitoring instrument codebook and guidelines). The instrument was used **each time** a client had a guidance session, although not all fields were completed at every session<sup>6</sup>, as some were relevant to first sessions only and others designed to collect exit data. Each client was assigned a unique identifier by the counsellor, allowing evaluators to link data for clients who participate in multiple sessions.

At the project proposal stage, each country set a target for the number of service users that their guidance programme would reach (see Table 3.1.).

**Table 3.1.** *Target and realised sample sizes, GOAL*

Partner country	Wave 1 target (number of clients)	Wave 1 achieved (number of clients)	Full evaluation target (number of clients)	Full evaluation achieved (number of clients)
Czech Republic	50	15	100	132
Flanders	100	148	200	418
Iceland	50	21	100	95
Lithuania	50	50	100	100
Netherlands	100	8	200	76
Slovenia	75	49	150	160

Taking the end of Wave 1 data collection as the project's mid-point, we can see that two countries reached (or exceeded) the Wave 1 **sample size target**: Flanders (148 service users) and Lithuania (50 service users). The remaining four countries did not reach the sample size target in the first wave of data collection: Slovenia reached 49 service users (target 75), Iceland 21 service users (target 50), and Czech Republic 15 (target 50). With only eight service users in the Wave 1 data set, the Netherlands was furthest from the target sample size (of 100 adults). These eight clients all came from one of the Netherlands's four interventions sites: no service user data were collected from one

<sup>6</sup> GOAL counselling in Flanders began in April 2015, well in advance of the counselling programmes in the other five countries; a negative consequence of this early start was that the programme began before the roll out of the data monitoring instrument (November 2015). As a result only very limited data are available on 90 of the Wave 1 clients. For all other Flanders clients, full data are available.

site, as, contrary to expectation, client intake at this site was negligible. The remaining two Dutch sites were recruited to the study too late to include clients in Wave 1.

By the end of the full evaluation, however, recruitment targets were reached (or exceeded) by all but two countries. The Czech Republic surpassed their target of 100 clients by 32 and Slovenia exceeded their target of 150 clients by 10. Lithuania fully reached their target of 100 clients. Flanders recruited 418 clients, more than twice their target of 200. The remaining two countries did not reach the full sample size. However, Iceland was only five clients short. As in Wave 1, the Netherlands was furthest from the target sample size (76 out of 200 adults). However, this figure primarily highlights one of the challenges of the “light touch” approach in this country. In terms of administering the Literacy Screener through partner organisations, the Dutch GOAL project far exceeded its expectations: 1,525 Literacy Screeners were administered across the four intervention sites (far more than the intended 400). However, data collection challenges at the partner organisations meant that client monitoring data was collected for only 76 clients. (These challenges are detailed in the Netherlands’ national report.)

### Client Satisfaction Survey

The **Client Satisfaction Survey** was designed to gather data from service users about their experiences of counselling services. As discussed in Chapter 2, satisfaction surveys can play an important role in programme evaluations. However, it must be recognised that whereas clients are undoubtedly experts in terms of judging their own level of satisfaction with a service, this does not make them experts at judging the quality of the service.

The client satisfaction survey instrument was a short, two-page, self-completion survey offered in either **paper or online** formats. It contained eight questions: two gathered demographic information on the client (age, gender); five focused on the counselling session, and one question asked clients to record if they received assistance in completing the survey<sup>7</sup>. There were small differences between the surveys offered in the six countries, reflecting the different contexts in which the guidance was offered and the different objectives of various programmes. Flanders, the Netherlands and Iceland opted for an online version of the client satisfaction survey and Lithuania, Slovenia and the Czech Republic collected the client satisfaction survey data through paper-based questionnaires. Generally speaking, countries opted for the format that national evaluators felt would be easier for their service users to complete.

In Flanders client satisfaction survey data were collected through an online tool, with the survey completed anonymously by clients immediately after their first session. In the Netherlands clients whose Literacy Screener indicated that they had low basic skills were asked to complete an online client satisfaction survey at the end of the intake session (if necessary with the client manager’s support). In Iceland the client satisfaction survey was administered using the online survey software, *qualtrics*. After the first or second interview clients were shown into a room separate from where the interview took place. They took the survey using an anonymous survey link. Assistance was optional

---

<sup>7</sup> It was anticipated that low literacy levels, or migrants’ low skills in the national language, might prevent some clients from completing the survey without assistance.

and given, if needed, by a staff member at the programme site but not from the counsellor that just had the session with the client. Once clients had completed the survey they submitted their answers which were then automatically registered and simultaneously made accessible to local evaluators. For evaluators in Iceland, an electronic survey represented a more effective way of reassuring clients about anonymity than a paper-based survey did.

In the Czech Republic clients completed the client satisfaction survey in paper version distributed at the end of each session. Questionnaires were printed for the clients, as it was felt that offering the survey in this format would be easiest for clients. All of the questionnaires were filled by the clients themselves, without any help. Data from the questionnaires were then entered into Excel. After each session in Lithuania each service user was asked to fill in a paper-based client survey questionnaire about their experience of the guidance. Completed client satisfaction questionnaires were collected in a sealed box. All answers were then entered into an Excel file by local evaluators. The Slovenian survey was also paper-based.

The total number of completed client satisfaction surveys is 804. In all countries, clients could choose to receive assistance in completing the survey (although not from the staff members they had been counselled by). Overall 20% of service users (162 out of 804) had some **help in completing the survey** (see Table 3.2.). No help was needed in Lithuania; help was needed in only a few cases in the Czech Republic and Slovenia. Around one-third of service users from Flanders and almost half from Iceland and the Netherlands had help to complete the client satisfaction survey.

**Table 3.2.** *Did anyone help you complete this survey?*

Country		Yes		No		Total	
		N	%	N	%	N	%
Country	<b>CZ</b>	3	2	132	98	<b>135</b>	<b>100</b>
	<b>FL</b>	72	30	165	70	<b>237</b>	<b>100</b>
	<b>IS</b>	44	48	47	52	<b>91</b>	<b>100</b>
	<b>LT</b>	-	-	100	100	<b>100</b>	<b>100</b>
	<b>NL</b>	32	43	43	57	<b>75</b>	<b>100</b>
	<b>SL</b>	11	7	155	93	<b>166</b>	<b>100</b>
	<b>Total</b>	<b>162</b>	<b>20</b>	<b>642</b>	<b>80</b>	<b>804</b>	<b>100</b>

In Lithuania the number of completed client satisfaction surveys was equal to the number of clients; in other words, each service user completed the survey once, generally after completing their first counselling session (see Table 3.3.). In both Iceland and the Netherlands, all but one client completed the survey. In the Czech Republic and Slovenia there were a (slightly) greater number of surveys than clients because some of the clients who had more than one session filled in the satisfaction survey after each session rather than just their first one. A far greater shortfall in completed surveys can be observed in Flanders, where 58% of service-users filled in the satisfaction survey. This shortfall can partly be explained by the fact that, as with the data monitoring, some clients in Flanders were engaged in the guidance programme before the satisfaction survey went into the field. Additionally,

local evaluators indicated that this poor response, particularly in Ghent (de Stap), was likely related to some clients' language difficulties.

**Table 3.3.** Satisfaction survey sample, N of respondents by country

	Service users	Satisfaction survey respondents	
	N	N	% (out of service users)
<b>CZ</b>	132	135	102
<b>FL</b>	418	241	58
<b>IS</b>	95	94	99
<b>LT</b>	100	100	100
<b>NL</b>	76	75	99
<b>SI</b>	160	166	104
<b>Total</b>	<b>981</b>	<b>811</b>	<b>83</b>

### Client follow-up survey

Beginning in January 2017, a **mixed methods follow-up survey** was conducted with 148 clients, either face-to-face or via phone, depending on the country. This survey sought to collect longitudinal outcome data from clients approximately two to four months after they left GOAL, in order to provide evidence on programme effects. The survey also collected quantitative and qualitative data on clients' perspectives on the programme and its impacts on their lives. As we can see from Table 3.4., three of the countries reached their target of 30 clients for the follow-up survey. As discussed in greater detail in the national reports, recruitment for this follow-up survey typically proved challenging: relatively few clients were willing to give permission to be contacted after leaving GOAL, and amongst those who did agree to be contacted a significant proportion either were not contactable (e.g. because of a changed phone number) or refused to participate.

Though the follow-up survey was conducted in as rigorous a manner as possible, any conclusions drawn from it must be considered tentative: the limited time period of the evaluation means that only short-term outcomes could be assessed. Furthermore, participation in the follow-up survey was voluntary and thus non-representative; therefore it is not possible to generalise from the survey findings to the broader group of GOAL clients. That being said, the survey may provide valuable insights into the experiences and outcomes of particular clients or subgroups of clients. As discussed in Section 2.1, this evidence may thus contribute to programme theory regarding how to meet the needs of such clients.

As Table 3.4. demonstrates, the overall number of clients is quite uneven, both when taking account of study waves and when comparing across countries. This needs to be taken into account when analysing the data available to distinguish between overall patterns and some more country specific characteristics and trends.

**Table 3.4.** Data Collection by country, showing Waves 1 and 2 (W1, W2) and Total (T)

Method	CZ			FL			IS			LT			NL			SI		
	W 1	W 2	T	W 1	W 2	T	W 1	W 2	T	W 1	W 2	T	W 1	W 2	T	W 1	W 2	T
<b>Monitoring data</b>	15	117	132	91	327	418	21	74	95	50	50	100	8	68	76	49	111	160
<b>Client satisfaction survey</b>	18	121	139	79	162	241	21	73	94	50	50	100	7	68	75	49	117	166
<b>Follow-up survey</b>	-	18*	18	-	30	30	-	22	22	-	31	31	-	14	14	-	33	33

\*Information from only 6 clients was used in this evaluation report as some of the CZ data collection took place too late to include the additional clients in the cross-country dataset.

### 3.3 Qualitative sample and data collection

Qualitative data was collected at two different stages over the life of the programme. The first stage (Wave 1) of qualitative data collection took place in Spring 2016. The second stage of qualitative data collection took place in Spring 2017. By collecting such data in two waves rather than only one, the evaluation is able to provide a longitudinal focus on issues explored through the qualitative analysis.

During each data collection wave, **semi-structured qualitative interviews and/or focus groups** were conducted with a range of programme stakeholders. In the first stage of qualitative data collection, four topic guides were developed by IOE to assist local evaluators in Wave 1 data gathering and to ensure consistency across the programme locations: **1) Programme Staff; 2) Programme Partners; 3) Policy Actors; and 4) Service Users**. A fifth Topic Guide, used in Iceland only, combined questions for Programme Partners and Policy Actors.

Similar topic guides were developed for the **second wave of qualitative data collection**. However, the development of these later topic guides was led primarily by local evaluators in each country, in consultation with IOE. The second wave of topic guides was somewhat more targeted in terms of the issues that were focused on: after qualitative data from Wave 1 had been analysed, key issues requiring further exploration or understanding were highlighted. These issues then formed the basis for Wave 2 topic guide development.

In Wave 2, interviews with service users were conducted as part of the mixed methods client Follow-up Survey (see section 3.2).

**Table 3.5.** *Qualitative data sample by country, Waves 1 and 2*

	Wave	Clients (N)	Programme Staff (N)	Programme partners (N)	Policy actors (N)
<b>CZ</b>	1	2	2	0	3
	2	18	2	2	2
<b>FL</b>	1	4	6	24	5
	2	30	8	8	8
<b>IS</b>	1	2	4	4	2
	2	22	4	4	2
<b>LT</b>	1	8	4	4	6
	2	31	4	4	6
<b>NL</b>	1	0	4	7	5
	2	14	4	4	4
<b>SI</b>	1	2	8	8	6
	2	33	4	4	8
<b>Total</b>		<b>166</b>	<b>54</b>	<b>73</b>	<b>57</b>

*N = number of interviewees / focus group participants.*

*NB: Clients, programme staff, partners and policy actors may have been interviewed in Waves 1 and 2, thus totals for these groups do not represent the total number of individuals interviewed.*

### 3.4 Programme staff: quantitative background survey

In Wave 1, a short quantitative survey was administered to all **GOAL programme counsellors** (not only those participating in qualitative interviews and focus groups) to gather some basic data on their educational background, their current employment, and their professional development and training.

This programme staff background survey includes data on 29 members of staff from five out of the six countries. In the Netherlands, no quantitative information was collected from programme staff as evaluators chose not to burden participating organisations further, because of the complex situation (for example, in the case of the two participating prisons) and because staff data were considered as less relevant given the light touch nature of the Dutch intervention. Data were received about four staff members in each of the three project countries (see Table 3.6.): the Czech Republic, Iceland and Lithuania. Slovenia provided data on 10 programme staff and Flanders on seven.

**Table 3.6.** *Staff background survey sample, by country*

		N	%
<b>Country</b>	<b>CZ</b>	4	14
	<b>FL</b>	7	24
	<b>IS</b>	4	14
	<b>LT</b>	4	14
	<b>SI</b>	10	34
	<b>Total</b>	<b>29</b>	<b>100</b>

### 3.5 Data analysis

In this **mixed methods evaluation**, a number of analytical approaches were used. Quantitative data were analysed using mainly **descriptive statistics**, e.g. frequencies, averages, group comparisons and cross-tabulations. Where appropriate, in particular with regard to outcomes (see Chapter 11), some inferential statistics in the form of **bivariate analysis** and **regression models** were also used to explore the impact of the programmes while controlling for individual client characteristics.

Raw data from national monitoring data collection, client satisfaction surveys and programme staff surveys were merged into one database and used for all the analysis in cross country reports. Where some variables needed **coding and banding** this was done from the raw data rather than using existing codes from the national datasets, in order to check the data and its consistency across all six countries.

Qualitative data from focus groups and interviews were analysed using mainly thematic analysis around the topic guides that was explicitly linked to the main aims and objectives of the project. In addition some typology analysis and group comparison analysis were also used.

### 3.6 Methodological challenges and study limitations

#### Differences in national programme models

The unique nature of the GOAL intervention in the Netherlands presented challenges for the cross-country evaluation, and for the reporting of this evaluation. In the current report, there are a number of instances in which messages from or about the other five countries are not relevant to the Netherlands, and vice versa. As much as possible, that is made clear in our reporting.

More generally, differences in national target groups and intervention types across the other five GOAL countries make it difficult and perhaps misguided to attempt to compare programme outcomes. Due to the large number of differences in programme target groups, resources and policy environments, it was generally not possible to isolate quantitative variables and point to these as key factors in shaping client outcomes.

#### Data collection

There were a number of initial challenges associated with the collection of robust **monitoring data**. Because **programme rollout typically took longer than expected**, many clients were still in an early stage of counselling during Wave 1 of the evaluation, and as a result little outcome data was available for the interim (Wave 1) report. For example, all the clients in Iceland and most of those in Slovenia were still in the guidance process at the Wave 1 cut-off date, meaning that it was not feasible to measure even relatively short-term programme impacts. In some countries the programme model allowed for one guidance session only, meaning there was no true exit data.

Evaluators from Flanders reported that staff members at one of the programme sites needed time to build up experience and familiarity with the data monitoring system. (Existing registration systems at

de Stap were not suitable for the collection of GOAL data.) Flemish counsellors felt that the data registration system required too much time during the session, so both de Stap and de Leerwinkel developed their own forms to make notes during the sessions. They then entered the data in the electronic system after the session. This placed additional burdens on counsellors. Initial resistance to the volume of monitoring data teams were required to collect by IOE was not unique to Flanders. Although fears about the onerous of this system soon proved to be unfounded, they draw attention to the fact that under modern managerial and workload efficiency pressures, it is **difficult to add an extra task into the counsellor's workload**, especially where it is unclear to the organisation how this additional work will bring them direct benefits.

With regard to the **Client Satisfaction Survey**, several countries reported that it was a challenge for some service users, especially those with poor basic skills, to evaluate their experiences. Lithuanian evaluators reported that clients with low levels of education found it challenging to summarise, analyse and reflect on their experiences.

As discussed above (Section 3.2), the **Follow-up Survey** presented a number of data collection challenges with regard to recruiting participants. Furthermore, the voluntary, self-selecting nature of participation in this survey greatly increases the likelihood of selection bias; thus these survey results must be seen as indicative rather than generalisable. However, there is evidence of consistency between findings of the follow-up survey and the client monitoring data. As part of the monitoring data, clients who had more than one guidance episode were asked during their last session (N=432) if they had taken the steps they hoped to. Two thirds (66%) of these clients agreed fully that they had taken the planned steps and 23% partially agreed. Only 12% of clients reported that they had not taken the steps they hoped to at the point of their last guidance session. Despite these results covering a much broader (and less self-selecting) range of GOAL clients than the follow-up survey results, the findings from the two instruments were very similar, suggesting that follow-up survey respondents are not necessarily unrepresentative of GOAL clients who had more than one guidance session. However, when generalising it is important to err on the side of conservatism and thus to avoid generalising from the follow-up survey to the broader GOAL client group.

### Challenges associated with measuring education and employment outcomes

A key implication of the fragmented, complex nature of adult education (see Chapter 12) is the difficulty of tracking clients over meaningful periods of time following programme exit. The lack of data sharing across educational institutions and policy sectors makes the **longitudinal tracking of outcomes** from interventions such as GOAL methodologically problematic, in large part because of the tremendous resource investment that would be required both to: a) track programme participants for sufficient time after programme participation, and b) establish and maintain matched comparison groups. In the absence of sufficient – and sufficiently rigorous – longitudinal tracking of programme participants, evaluation assessments of programme impacts are merely indicative, and make it difficult for programme developers and policy makers alike to assess the true impact and thus value of the intervention.

This lack of longitudinal data is an unfortunate characteristic of almost all evaluations in the adult education field: in contrast to interventions focused on school children, it is extremely difficult to track adults and their outcomes after they leave interventions. This challenge would be greatly alleviated by the use of and access to harmonised data across policy domains, which would allow, for example, the tracking of GOAL clients' future educational and employment outcomes. However, data sharing is a contentious topic, and national policies on data protection vary. In the GOAL project, data policy influenced what information could and could not be collected by the national evaluation teams, and what information could be shared within countries. Privacy norms and historical lack of joined-up policy making mean that shared data were not available in the GOAL countries. This lack of shared longitudinal data has implications for counselling services seeking to develop their programmes on the basis of what has been learned, or attempting to construct referral arrangements to other organisations. Data-sharing privileges and tools would support partnership working and the accurate measurement of longer term programme outcomes.