Reshaping Attention and Inclusion Strategies for Distinctively vulnerable people among the forcibly displaced

www.raisd-h2020.eu

Evaluation Criteria

Actor–oriented and integrated evaluation

Deliverable D7.3

Antti Kivijärvi
University of Helsinki, Unioninkatu 37 (PL 54), 00014, Helsinki, Finland

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 822688.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 822688.

Document Information

Grant Agreement #:	822688
Project Title:	Reshaping Attention and Inclusion Strategies for Distinctively vulnerable people among the forcibly displaced
Project Acronym:	RAISD
Project Start Date:	1st February, 2019
Related work package:	WP7 Cross analysis
Deliverable	D7.3 Evaluation criteria: actor-oriented and integrated evaluation - preliminary version
Related task(s):	T7.3 Definition of evaluation criteria
Lead Organisation:	University of Helsinki
Submission date:	31/01/2021
Dissemination Level:	Public

History

Date	Submitted by	Reviewed by	Version (Notes)
13/02/2020	Antti Kivijärvi (UH)	Clara Guilló; Liisa Hänninen (UCM)	First version (draft)
29/02/2020	Antti Kivijärvi (UH)	All partners	First version (final)
29/04/2020	Antti Kivijärvi (UH)	Liisa Hanninen; Clara Guilló (UCM)	Second version
03/06/2020	Antti Kivijärvi (UH)	All partners	Third version
31/01/2021	Antti Kivijärvi (UH)	All partners	Public version

Coordinator contact: Dr. Rubén Fuentes-Fernández | Universidad Complutense de Madrid | Calle del Profesor José García Santesmases, 9. Ciudad Universitaria 28040 MADRID, Spain.
t: +34/ 91 394 7548 | e: rfuentes@ucm.es | w: www.ucm.es , grasia.fdi.ucm.es

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 822688.

Table of content

List of Tables

List of Figures

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 822688.

RAISD Glossary

ARU Action Research Unit

EC European Commission

FDP Forcibly Displaced People / Person

HVG Highly Vulnerable Group

LIU Lebanese International University

LGTBIQ Lesbian, Gay, Transgender, Bisexual, Intersex and Queer or Questioning

MOU Memorandum of Understanding

NGO Non-Governmental Organisation

PRSF Psychosocial Refugee Support Forum

RAISD Reshaping Attention and Inclusion Strategies for Distinctively vulnerable people among the forcibly displaced

RCT Randomized Controlled Trial

RRI Responsible Research and Innovation

RSS Refugee Support System

SMART Specific, Measurable, Attainable, Relevant and Time-based

TAIS Tailored Attention and Inclusion Strategy

UNHCR United Nations High Commissioner for Refugees

VC Vulnerability Context

VG Vulnerable Group

WP Work Package

YU Yarmouth University

Coordinator contact: Dr. Rubén Fuentes-Fernández | Universidad Complutense de Madrid | Calle del Profesor José García Santesmases, 9. Ciudad Universitaria 28040 MADRID, Spain.
t: +34/ 91 394 7548 | e: rfuentes@ucm.es | w: www.ucm.es , grasia.fdi.ucm.es

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 822688.

About RAISD	
Call (part) identifier	*H2020-SC6-MIGRATION-2018*
Topic	*MIGRATION-08-2018* *Addressing the challenge of forced displacement*
Fixed EC Keywords	*Globalisation, migration, interethnic relations*

Forced displacement crises overcome societies and institutions all over the world. Pushed by the urgencies rather than events, solutions are frequently reactive, partial, and disregard some groups. The project 'Reshaping Attention and Inclusion Strategies for Distinctively vulnerable people among the forcibly displaced' (RAISD) aims at identifying highly Vulnerable Groups (VG) among these forcibly displaced people, analysing their specific needs, and finding suitable practices to address them. The concept of 'vulnerability context' considers the interplay between the features of these persons and their hosting communities, their interactions and experiences, and how different solutions for attention and inclusion affect them. As a result of this work, a methodology to carry out these studies will be developed. These goals are aligned with the call. They pursue characterizing these migrations and developing suitable aid strategies for them. The Responsible Research and Innovation (RRI) frames the project. It proposes that all actors (including civil society) co-design actions, transversely integrates the gender perspective, and supports sustainability. Our research strategy will be based on methodological triangulation (i.e. the combined application of several methodologies). We will implement it through a specific participatory action research approach to fulfil the aim of undertaking advocacy-focused research, grounded in human rights and socio-ecological models. The team will work as a network of units in countries along migration routes. The units will promote the VG people' involvement, so they can speak with their own voices, gather information, and test practices. Work will rely on a tight integration of Social and Computer Sciences research. Automated learning and data mining will help to provide evidence-based recommendations, reducing a priori biases. A software tool will support collaboration, continuing previous H2020- funded RRI work.

Coordinator contact:
Dr. Rubén Fuentes-Fernández | Universidad Complutense de Madrid | Calle Profesor José García Santesmases, 9. Ciudad Universitaria 28040 MADRID, Spain.
t: +34/91 394 7548 | email: rfuentes@ucm.es | web site: www.ucm.es grasia.fdi.ucm.es

Executive Summary

In the RAISD-project, context-sensitive Tailored Attention and Inclusion Strategies (TAIS) are developed in order to alleviate the vulnerabilities of forcibly displaced persons (FDPs) living in different environments. To determine the feasibility of TAIS and to formulate evidence-based policy recommendations, an evaluation scheme is needed. Thus, the aim of this deliverable is two-fold.

First, a conceptual framework of evaluation is set. In RAISD, a realistic tradition of evaluation science will be endorsed. This means respecting the variety of stakeholder and end-user perspectives while acknowledging the different settings in which TAIS will be implemented – studying 'what works for whom in which contexts'.

Second, concrete evaluation criteria will be determined. In line with the RAISD framework, the deliverable will start from actor-oriented criteria of various stakeholders in different contexts.

Finally, based on these criteria, an integrated evaluation criteria and strategy will be defined. The deliverable will benefit from the works of work packages four and five. WP4 provides information on various vulnerability contexts and their specific features and needs of FDPs. WP5 provides information on partners' actor-oriented evaluation criteria and TAIS to be piloted. This report is a preliminary version of the deliverable 7.4 constituting the final version of the RAISD evaluation criteria and results.

Coordinator contact: Dr. Rubén Fuentes-Fernández | Universidad Complutense de Madrid | Calle del Profesor José García Santesmases, 9. Ciudad Universitaria 28040 MADRID, Spain.
t: +34/ 91 394 7548 | e: rfuentes@ucm.es | w: www.ucm.es , grasia.fdi.ucm.es

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 822688.

1 Introduction

In addition to producing knowledge on forced displacement and vulnerability contexts along the route from Middle East to Northern Europe, the essential aim of the RAISD project is to develop attention and inclusion practices tailored for refugee populations in vulnerable positions. Further, to assess the feasibility of these practices, methodological work will be constructed to establish common principles and tools of evaluation. In this deliverable, the focus is on the latter: evaluation of the Tailored Attention and Inclusion Strategies (TAIS) in different Vulnerability Contexts (VCs).

The aim of this paper is to provide a comprehensive idea about the evaluation methodology in RAISD. Like all RAISD activities, also the evaluation methodology has been developed from a participatory perspective. Consequently, the foundation of the methodology has been the evaluation criteria defined by various stakeholders in different contexts. These criteria are called actor-oriented evaluation criteria. Out of a diverse set of actor-oriented criteria, the task has been to determine integrated evaluation criteria while providing a common evaluation tool for RAISD project together with consortium partners.

In the following chapters five themes will be covered. First, a rough outline of the principles of evaluation in RAISD will be set. This task includes deciphering the notion of evaluation, presenting the contexts in which the evaluation will take place, showing where we stand in the large field of evaluation research and describing some key tools that can be used in the craft of evaluation.

Second chapter is dedicated to reviewing studies in which small-scale programs, services or interventions for refugee populations or forcibly displaced people have been evaluated. The glimpse in the literature provides an idea about the methods and shortages of evaluation studies in the context of forced displacement. Third chapter is based on contributions of each partner and their local stakeholders. In order to develop a reasonable evaluation scheme, it is necessary to know what types of interventions will be evaluated and how they are valued in each context.

In the key chapter (4) of this deliverable the 'integrated' evaluation criteria for the whole consortium is set. The integrated criteria will be derived from actor-oriented criteria and is adjusted to fit the theoretical and epistemological foundations of the evaluation framework. These foundations are based on arguments of realist evaluation. Realism refers to a ground between positivism and constructivism. While positivist evaluators often try to determine whether interventions work or not, constructivist evaluators ponder how are programs perceived among stakeholders. For a realist evaluator, the key slogan is 'what works for whom in which contexts'. Thus, for a realist evaluator, interventions are set in their discursive and material contexts and scrutinized among different stakeholders with mixed methods (contrary to positivist evaluation) with an aim to find mechanisms explaining changes in the real world of people (contrary to constructivist evaluation).

Finally, the evaluation methodology in RAISD is set to be congruent with the overall principles of the project: respecting diversity and principles of Responsible research and innovation (RRI) (Owen et al., 2012). Therefore, as with planning the TAIS and implementing them, it is of essential importance to closely co-operate with different stakeholders while setting the evaluation criteria for them as well. Moreover, as in all phases in the project, while determining the evaluation criteria, it is necessary to ensure the diversity among the stakeholders primarily in terms of gender but also in respects of several social categories such as ethnicity, sexuality and religion. Most

Coordinator contact: Dr. Rubén Fuentes-Fernández | Universidad Complutense de Madrid | Calle del Profesor José García Santesmases, 9. Ciudad Universitaria 28040 MADRID, Spain.
t: +34/ 91 394 7548 | e: rfuentes@ucm.es | w: www.ucm.es , grasia.fdi.ucm.es

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 822688.

importantly, including the FDP perspectives and co-designing various project activities with them is a necessity. Realist tradition provides a solid ground for promoting RRI and respecting diversity.

1.1 What is evaluation?

According to one renowned author in the context of evaluation studies, Evert Vedung, evaluation refers to 'careful retrospective assessment of interventions, their organization, content, implementation, outputs, or outcomes, which is intended to play a role in future decision situations.' (Vedung, 2010). In short, evaluation means defining 'the value of certain project, program, intervention or service' (Robson, 2000, pp. 24). Thus, evaluation inevitably includes political aspects and include at least some level of researcher engagement (ibid, pp. 25); the core idea is to evaluate interventions designed to make desired changes to happen.

Consequently, to deploy an old academic dichotomy, evaluation studies falls into the category of 'applied research' instead of a sometimes more distinguished 'basic research'. Conducting applied research requires abilities in doing field work, interacting with various stakeholders, an in-depth understanding of an empirical phenomenon under scrutiny and the contexts it is embedded in, and a capacity to compose policy relevant research settings. The evaluator is thus deeply bound up with the complexities, ambiguities and messiness of everyday life (e.g. Pawson, 2013, pp. 33-46).

In real life contexts outside laboratory conditions, conducting a reliable evaluation is never simple. First, changes are hard to accomplish (e.g. Stange & Phillips, 2007). A countless number of factors and institutional or structural conditions might mitigate the effects of even the most impressive program. A good evaluator can, at the very least, present informed predictions in which types of conditions certain interventions either succeed or fail. Second, if changes do happen, it might be even more difficult to indicate what it is in the program that have caused the changes. In other words, an evaluator is often expected to reveal causal relations between two or more variables or change provoking mechanisms (e.g. Webster & Sell, 2007). Third, after dealing with the above difficulties, an evaluator often faces an even trickier question about the 'scalability' of an intervention – whether and how certain practice can be reworked in or adjusted to other contexts. The unpleasant fact seems to be that good practices can turn into average or even bad ones while implementing it in another context. When contexts change, also the people and their ways and motivations to implement the program tend to change. Thus, in a real and complex world evaluations are always fallible. The evaluator finishes with tentative and context-bound answers.

In the RAISD-project and in this deliverable evaluation means defining the value of piloted activities, or TAIS, in their particular contexts. RAISD evaluation do not claim to reveal universal causalities but the aim is to, in line with the realist tradition, to explore some of the context-specific mechanisms promoting or preventing changes when trying to alleviate the vulnerabilities of specific refugee populations. This is done by having a thorough look into the 'black box' of each pilot with mixed and follow-up datasets and in close co-operation with a diverse set of stakeholders who are familiar with each context.

Finally, in RAISD, evaluation does not mean plainly investigating mechanisms and possible benefits of certain interventions. RAISD evaluation criteria are formulated by honouring the principles of sound ethical practice and reminders of RRI. In practice, this means involving stakeholders in different positions and including various forms of diversity in a process of co-creating the evaluation criteria.

Coordinator contact: Dr. Rubén Fuentes-Fernández | Universidad Complutense de Madrid | Calle del Profesor José García Santesmases, 9. Ciudad Universitaria 28040 MADRID, Spain.
t: +34/ 91 394 7548 | e: rfuentes@ucm.es | w: www.ucm.es , grasia.fdi.ucm.es

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 822688.

1.2 Pilots and their evaluation in RAISD

The number and backgrounds of FDPs and their living environments and vulnerability contexts differ greatly between the countries in the RAISD-consortium. The neighbouring countries of crises-ridden areas such as Jordan, Lebanon and Turkey host millions of FDPs while the number of refugees and asylum seekers in the 'fortress Europe' is significantly lower, particularly in some of its Eastern and Northern regions. In addition to differences in numbers, there are probably differences in characteristics as well. For decades, making the highly expensive, burdensome and dangerous journey to Europe, have required resources that only few have (e.g. Juntunen, 2016). Consequently, the poorest and physically weakest are often those who are not able to make the journeys while the bulk of asylum seekers in Europe are young men from families with at least some financial resources.

The most important differences within the RAISD consortium partner countries, however, might be contextual. Societal and cultural structures are quite different in prosperous Nordic states with strong public sectors and regulation compared to, for instance, Mediterranean countries and Middle East (Esping-Andersen, 1990; Fangen et al., 2012). With all these huge differences in mind, it should not come as a surprise that the pilots conducted will differ as well. TAIS probably need to be different, for instance, for asylum seekers in Finland than for undocumented Syrians in Turkey (for a detailed account on contextual differences, see deliverable 4.1).

To recognise contextual differences, all project work will follow an action research approach (e.g. Stringer, 2013). This means highlighting bottom-up and local knowledge acquired from various stakeholders and FDPs in order to plan best possible pilots and criteria to evaluate them. In practice, each project partner has established their local Action Research Units (ARUs), where researchers will co-operate in the field with various stakeholders and FDPs in their hosting communities. This approach will help to gain a deeper understanding of the contextual factors influencing the pilots, as well as the motivations and goals of different stakeholders.

As a result, the planned TAIS are quite different from each other. In Spain, the aim is to promote the inclusion of female refugees from sub-Saharan origins by enhancing their entrepreneurial skills. In Italy, the educational trajectories of women from refugee backgrounds will be supported with a help of a new digital tool allowing more individual agency. In Hungary, the pilot will target more professional audiences while the aim is to create a monitoring toolkit for social workers to recognise refugee vulnerabilities and find the most suitable services for them. Finnish two TAISs are focused on the national reception system and alleviating the vulnerabilities of two asylum seeking groups: young single men and parents of small children. Multilingual online forum aims to foster language acquisition and social inclusion of asylum-seeking young men while the goal of developing childcare activities in reception centres is to improve the living conditions of both parents and their children.

Outside European Union, in Turkey, Lebanon and Jordan, the planned TAIS take their unique and context sensitive forms as well. In Turkey, a series of trainings and a monitoring tool are provided for professionals to promote the rights of mainly Syrian female refugees in their local areas. In Lebanon, an online health awareness campaign will be delivered. The campaign is geared to engage in awareness, prevention and treatment of COVID-19. The aim is to empower civil society agencies to facilitate and support the teaching of Syrian, Palestinian and other displaced people in refugee camps. Finally, in Jordan, methods of delivering psycho-social support for refugee populations are developed. The main goal of the program is to reduce asylum seekers' need for remedial and clinical support.

Coordinator contact: Dr. Rubén Fuentes-Fernández | Universidad Complutense de Madrid | Calle del Profesor José García Santesmases, 9. Ciudad Universitaria 28040 MADRID, Spain.
t: +34/ 91 394 7548 | e: rfuentes@ucm.es | w: www.ucm.es , grasia.fdi.ucm.es
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 822688.

1.3 Fallible realism in real world settings

One can find a countless number of evaluation cookbooks, a huge number of distinct evaluation methods and even a few paradigms of evaluation studies. In the literature on evaluation studies, a couple of illustrative distinctions can be observed. First, on one hand, there is positivist and empiricist stances embracing natural sciences, experimental trials and quantitative methods. On the other hand, one can see evaluation studies emerging from relativist and constructivist thinking while giving more weight on social sciences and qualitative methods. Second, and strongly related to the previous distinction, different sets of data are appreciated while conducting evaluations. Others prefer fixed top-down settings and objective data which can be produced independent of the stakeholders (e.g. register data), while others embrace more flexible bottom-up models and aim to produce more subjective knowledge with the program participants and stakeholders. (e.g. Vedung, 2012; Robson & McCartan, 2016, pp. 5-6.)

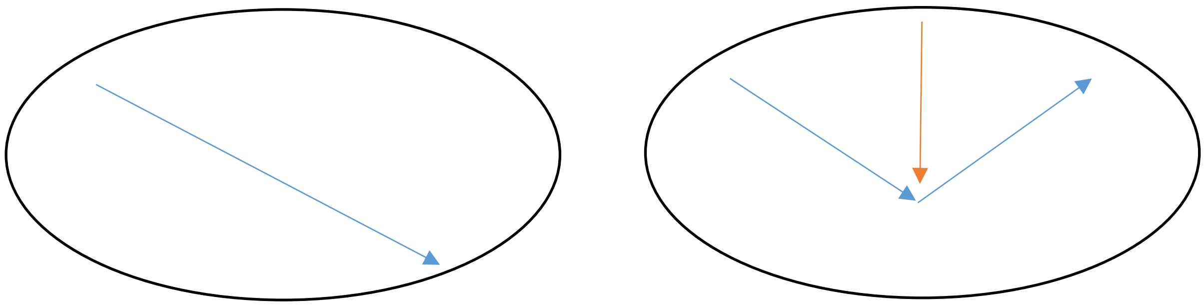
While evaluating the RAISD pilots, the overarching aim is to find a middle ground between the above (and often artificial) juxtapositions. The inescapable starting points in RAISD are that evaluated pilots will be implemented in differing contexts, with various aims and with different populations and stakeholders. Common for most of the pilots is that they will be relatively small-scale and local exercises (in contrast to state-level programs) and they will be implemented in real world settings with actual people (in contrast to controlled laboratory conditions). Thus, the aim is to rule out both purely positivist and relativist methods. The burden of the former is context and complexity blindness and giving little value to stakeholder and end-user knowledge while the burden of the latter is getting lost in thick interpretative haze (e.g. Vedung, 2012; Pawson, 2013).

Consequently, as a loose paradigm of evaluation, the RAISD consortium resorts to realism. The starting point of realism is that the reality is 'out there', largely independent of the human mind. Even though the 'reality' is not beyond dispute, and to a great extent a social and historical product, people face it in their everyday lives by getting things wrong and by crashing into objects (e.g. Sayer, 2000, pp. 2). However, contrary to most positivists and empiricists, according the scientific realism the truth is not directly observable. In the spirit of Roy Bhaskar and Karl Popper, the aim of scientific realism is to make sense and theorise observable regularities and their exceptions in certain contexts during certain times, not to claim universal truths.

The above statement means that instead of trying to reveal directly observable 'successionist' causality' (A causes B because the occurrence of A is followed by the occurrence of B), the aim is to explore 'generative causation' (B follows A because of operation of certain mechanisms). The key word in realistic evaluation is thus mechanisms (see e.g. Pawson & Tilley, 2004, pp. 6). Mechanisms describe what in programs bring effects (figure 1). However, while mechanisms are often hidden (not directly observable), they are not the results of certain program as such but possible responses of various stakeholders. Programs thus might activate certain mechanisms. (E.g. Pawson & Tilley, 1997, pp. 65; Robson, 2000, pp. 115-116; Astbury and Leeuw, 2010). Realist slogan says that 'interventions do not work, it is the interpretations of their subjects that produce results' (Pawson, 2013, pp. 34). Consequently, people are seen as actors in their real-life settings, not passive recipients of some exterior interventions.

Coordinator contact: Dr. Rubén Fuentes-Fernández | Universidad Complutense de Madrid | Calle del Profesor José García Santesmases, 9. Ciudad Universitaria 28040 MADRID, Spain.
t: +34/ 91 394 7548 | e: rfuentes@ucm.es | w: www.ucm.es , grasia.fdi.ucm.es

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 822688.



Context =
Outcome =
Mechanism =

Figure 1. Mechanisms and generative causation (a revision of a figure by Robson & McCartan, 2017, pp. 33)

All this means that an evaluator needs to have a substantial knowledge on the context and the phenomenon under investigation and possible mechanisms that might play a role (either promoting of blocking certain effects). In real world surroundings, contrary to aims in several positivist and experimental settings, mechanisms and their contexts cannot be controlled. Real world research often takes 'place in someone else's territory' (Robson & McCartan, 2016, pp. 1). In settings such as schools, welfare agencies and refugee camps, an evaluator 'have far from perfect control, are typically guests and not royalty, have to negotiate and not command, and often must compromise rather than get everything they would like' (Shadish, Cook & Campbell, 2002, pp. xix). Thus, instead of outside observer, an evaluator is often more like a co-implementer and a stakeholder.

A popular slogan for realist evaluators is 'what works for whom in which contexts' (Pawson, 2013, pp. xiii). A realist evaluator thus admits the limits of scientific enquiries – they are always fallible at least to some extent. Descriptions of generative causation (how mechanisms produce events) are thus best possible explanations based on limited empirical evidence. They cannot be taken as universal laws of (successionist) causality. For a realist evaluator (contrary to a positivist one), 'context is not unwelcome noise, not a confounding variable to be controlled for – but an integral part of a programme' (Pawson, 2013, pp. 36). Thus, the aim is not to exclude the effect of contexts or unknown factors, but rather to take them as objects of investigation (see e.g. Patton, 2014). To be able to reveal the 'black box' of programs, to determine what makes them either work or fail, an evaluator needs to resort, in addition to contextual knowledge, to mixed data-sets (e.g. Robson & McCartan, 2016, pp. 5-6; Lipscomb, 2008). Concretely, this might mean producing different types of data (e.g. quantitative and qualitative) with various stakeholders (e.g. service providers and end-users) in different phases of the intervention (before, during and after).

In the end, an obvious challenge of realist evaluations is the question of how to treat the contextual complexity in a systematic manner and be able to say something meaningful about actual mechanisms. According to Ray Pawson (2013, pp. 21-22), the key to success is to use the 'if-then proposition'; if the right processes operate in the right conditions then the intervention will prevail. The task is thus to explain how contexts, mechanisms and outcomes are tied together, not just to list them in an arbitrary manner (ibid., pp. 26-27). To illustrate this type

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 822688.

of thinking, a gunpowder analogy is often used. Gunpowder do not blow up when flame is applied. Only if it is dry and warm enough, then applying flame to gunpowder will generate a blowing reaction. Revealing these types of 'ifs' and 'thens' are the essential targets of realist evaluation (again, see figure 1).

1.4 Towards evaluation practice: categorizations and distinctions

For the purposes of RAISD-consortium, in addition to sketching theoretical foundations, it is necessary to present some practical tools of evaluation. However, since RAISD is a multidisciplinary and highly international consortium, the practice of evaluation cannot be too fixed and some leeway for each partner need to be allowed. Moreover, respecting the principles of RRI and participation of various stakeholders are likely to cause some differences between different partners and in their actor-oriented evaluation criteria.

The most obvious requirements in the practice of realist evaluation are having evaluators familiar with a given context and the need for mixed-methods and follow-up datasets. Thus, all pilots will be evaluated by local experts with local stakeholders and beneficiaries. Local ARUs (Action Research Units) are responsible of doing this. Moreover, the aim is to combine various datasets such as quantitative and qualitative, subjective and objective, while collecting them from a variety of stakeholders along the trajectory of the pilots.

In the literature on evaluation practice, one finds a time-based hierarchy of evaluation targets (e.g. Scriven, 1967; Robson, 2000). First, needs should be evaluated. In RAISD, the needs of FPDs have already been analysed quite thoroughly during earlier phases of the project (see D4.1 and D5.1) and this will be continued throughout the project trajectory. Second, the process of the pilot or program should be evaluated. Process evaluation includes questions such as does the program match the professional or institutional standards of the stakeholders or the needs of beneficiaries and what happens during the pilots and how can they be improved? Process evaluation often requires qualitative and subjective datasets. Questions need to be posed to those people who are involved with the program. Obviously, process evaluation can include quantitative and objective measures as well. For instance, indicators such as adherence to programs are often used to evaluate the acceptability of the intervention (e.g. Eldridge et al., 2016).

The third phase in evaluation practice is often the impact or outcome evaluation. The essential questions included in the outcome evaluation are: Does the program meet the goals that have been set? Does it provoke the desired changes? What kind of intended and unintended causes the program has? Consequently, outcome evaluation often includes grading the program based on its merits. Outcome evaluation are often carried out with quantitative methods to observe measurable changes in participants' behaviour or perceptions without excluding the possibility of retrospective qualitative inquiries. Fourth, in some evaluation studies, particularly in those conducted by economists, the (cost-)effectiveness of the program is calculated. In these cases, the relation between program benefits and costs are examined to decide whether the program should be recommended for policy makers. (see e.g. Robson, 2000, pp. 77-80).

Coordinator contact: Dr. Rubén Fuentes-Fernández | Universidad Complutense de Madrid | Calle del Profesor José García Santesmases, 9. Ciudad Universitaria 28040 MADRID, Spain.
t: +34/ 91 394 7548 | e: rfuentes@ucm.es | w: www.ucm.es , grasia.fdi.ucm.es
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 822688.

Table 1. Key phases of evaluation practice

Phase	Tasks of evaluator	Methods	Timing
Needs evaluation	Desk research and conducting new empirical data and analysing it	Literature, statistics, register data and collecting information from key stakeholders	While planning the program
Process evaluation	Determining what happens during the intervention, revealing mechanisms	Mostly qualitative: Engagement with those who implement or are involved with the intervention	During the program implementation
Outcome evaluation	Determining the level of change the intervention provokes, defining key measures/indicators	Mostly quantitative: baseline, interim and follow-up measures	Before, during and after the program implementation

The division between needs, process and outcome evaluation can be seen in going hand-in-hand with the 'context, mechanism and outcome configuration' presented in the framework of realist evaluation (see Pawson, 2013, pp. 21-27). In the realist discourse, contexts correspond to needs. Needs inevitably arise out of the interplay between human existence in certain contexts. Mechanisms, for one's part, can be considered processes triggered by the program or intervention. They can include processes such as co-operation, enabling new relationships, awareness raising or improved access to services. Finally, outcomes are outcomes in both terminologies, things ensuing after needs/contexts are probed with processes/mechanisms.

2 A Literature review: Evaluation studies in contexts of forced displacement and exile

In this chapter studies empirically evaluating programs and interventions for forcibly displaced people in real life settings will be reviewed. Due to these limitations, the number of studies matching the criteria is quite modest, being 14. The abundant number of various descriptive reports on 'good practices' (e.g. EASO, 2018; EWSI, 2020) and studies on, for instance, macro-level integration or citizenship indicators (e.g. Bauböck & Helbling, 2011; Beine et al., 2016; Helbling et al., 2017) are disregarded. Further, also the multitude of (often non-academic) evaluation reports by UNHCR on relatively large-scale refugee programs (see UNHCR 2020) are excluded from this review. Evaluation studies concerning interventions on unaccompanied minors are excluded as well since it is not the focus of the RAISD consortium. Finally, the small number of studies included in this review is explained by the fact that three of the reviewed studies are themselves reviews including dozens of studies. Consequently, the number of evaluation studies in particularly medical and psychological paradigms utilizing experimental methods is relatively large. However, the main focus of this review is on social scientific studies.

For this version of the deliverable, no systematic review has been made. Searches have been conducted in Google Scholar and in some of the key academic journals (Journal of Refugee Studies, Refuge Survey Quarterly, Journal of Ethnic and Migration studies and Journal of Immigrant and Refugee Studies) with keywords 'evaluation',

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 822688.

'intervention', 'pilot', 'effect' and 'feasibility'. With this method several studies are very likely missed but the result provides an overall picture about methods of evaluation in this particular context.

2.1 RCT as a 'golden standard' of evaluation

Evaluating an intervention or a program is most often approached from a positivist paradigm of examining the effectiveness and fulfilling the criteria for a randomized controlled trial (RCT). RCT settings are often labelled as the 'golden standard' of evaluation and in determining program effectiveness. Particularly the fields of medical sciences, psychology and economics are dominated by RCT studies while these disciplines, for instance, include a huge number of exclusive meta-studies aiming to summarize the empirical findings of studies deploying RCT methodologies. In short, RCT studies aim to utilize the basic logic of experimental research: conducting at least two-arm settings in which the only difference between the experimental and control arms is the application of certain intervention. Despite the consistent critique toward experimental research in particularly complex human conditions (e.g. Patton, 2014; Tilley & Pawson, 1997), RCT methods seem to dominate the evaluation research in refugee and FDP contexts as well.

Out of the 14 studies matching the criteria of this literature review, four deployed, to some extent, an RCT setting with either randomized study arms or quasi-experimental methods. Three of these studies were written by either psychologists or representatives of medical science. One paper was written by economists. Psychological and medical papers investigated the 'effectiveness', 'effects', 'impacts' or 'efficacy' of various refugee programs. The domination of experimental setting is highlighted by the fact that two of these studies were research reviews including only RCT studies as their sources of information. Altogether they included 37 individual studies (not analysed individually in this review). In any case, in both of the research reviews it is claimed that psychological, mostly therapeutic interventions have some positive effects on refugee populations in alleviating anxieties, depression and post-traumatic stress disorder. (Sayedhabibollah et al., 2019; Turrini et al., 2019) Consequently, it seems evident that psycho-medical disciplines and experimental methods dominate the evaluation literature in also refugee and FDP contexts.

The two other experimental or quasi-experimental investigations included one conducted by psychologists and one by economists. The former study aimed to determine the effectiveness of 'life-skills' training for Tibetan refugee adolescents in reducing their stress levels (Yankey & Biswas, 2012) and the latter investigated the impact of smart-phone technologies on 'social and economic integration' of refugees in US (Evans et al., 2019).

The domination of experimental methods is even further evidenced when looking the studies more closely. In addition to above four studies, five papers draw from a tradition of so called 'feasibility' or 'pilot studies' investigating the potential effects of certain programs and possibilities to expand the study into an RCT setting (on pilot and feasibility studies in more detail, see e.g. Tickle-Degnen, 2013; Eldredge et al., 2016). Conducting a feasibility study often means utilizing a mixed-method approach to determine the acceptability and viability of the intervention from a stakeholder and end-user perspectives, in addition to providing a further research agenda for an experimental setting (see e.g. Kivijärvi et al., 2019).

As with experimental settings, feasibility studies investigated are authored by either psychologists or medical scientists. The interventions evaluated include promoting mental health with various discussion groups and individual counselling (Huaibo et al., 2011), treatment of type 2 diabetes with a digital storytelling method

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 822688.

(Wieland et al. 2017), increasing language skills with a literacy program (Kupzyk et al., 2016), making health-care services more accessible with a tailored method of awareness raising (Piwowarczyk et al., 2013) and decreasing alcohol related harms with a training of refugee camp workers and tailored services (Ezard et al., 2010). Methods of the above studies are somewhat typical for feasibility studies. They include subjective enquiries in quantitative and qualitative form such as baseline and follow-up surveys and feedback discussions. Moreover, two studies include also objective measures, one on receptive vocabulary (Kupzyk et al., 2016) and the other utilizing medical records (Wieland et al., 2017). Common to all feasibility studies is their somewhat descriptive nature and lack of control groups.

2.2 Other evaluation strategies in refugee and FDP contexts

There seem to be only few studies not resorting to the positivist discourse of experimental research. The obvious shortage of 'alternative' ways of evaluating interventions even in highly complex and varied FDP contexts might reflect the criteria of scientific journals and other academic platforms of publication. In any case, the result can be considered as unfortunate since experimental methods may not always be the only and best solution when aiming to evaluate interventions in changing and diverging contexts outside laboratory conditions.

It might not come as a surprise that all articles falling outside the experimental and feasibility tradition are written by social scientists. At the same time, however, it is difficult to determine the scientific grounds of the papers while they include various (mostly qualitative) methods. In some papers, it is claimed that the effects, impacts or even effectiveness is scrutinized which is quite untypical outside experimental and positivist literature. The above bundle of studies investigates the impact of microcredit program in Uganda (Jacobsen et al., 2006), the effect of Frontex border guard training (Horii, 2012) and one paper is a review of three studies evaluating the support systems for victims of human trafficking (Davy, 2015).

In the study on microcredit program, the 'impact' of the program in improving the living conditions and increasing the economic security of FDPs is evaluated with a great multitude of data sets. Authors utilize 'a household economic portfolio model' to investigate household economic activities, baseline and follow-up surveys with a comparison (not a control) group and qualitative interviews of FDPs and other stakeholders. (Jacobsen et al. 2006.) It is likely that this type of mixed-methods enquiry provides a lot of contextual information and enable to answer a wide range of questions concerning the usefulness of the program. The paper evaluating the effects of Frontex border guard training utilizes more simple data sets. While the aim of the author is to study the professional socialization and convergence of border guards as a result of training, she uses qualitative interviews of border guards and other stakeholders. (Horii, 2012.)

As for the review of evaluations on human trafficking programs, all three studies are located in the US. All three studies evaluated different things in somewhat similar programs. The first study investigated the outcomes of the program in the form of promoting the wellbeing of the clients by utilizing a variety of methods: logic model and retrospective interviews of key stakeholders. The study included no direct data from end-users (trafficked persons) but analysed documents created by workers and clients together (Potocky, 2010). Instead of outcomes, the second study focused primarily on processes of service provision (some questions on outcomes were presented as well). In practice, this meant investigating the planning, implementation, accessibility, delivery and sustainability of the services. The author used qualitative interviews of various stakeholders and victims/survivors of human trafficking. (Caliber, 2007.) In the third study, the accessibility of the services for end-users and

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 822688.

availability of training for professionals were evaluated. The evaluation focused on professional stakeholders and victims of human trafficking. The authors used several instruments for their evaluation, including questionnaires with opened-ended questions for professionals and victims, and a client satisfaction survey. (Busch et al., 2007.)

The remaining two studies are qualitative enquiries claiming to be either a case study (Cambridge & Williams, 2004) or 'a reflection' (Kaiser, 2004). The former is investigating the 'potential for advocacy' in social services for asylum seekers and refugees in UK. The aims of the evaluation are not stated in explicit manners but developing advocacy in the context of social services is probably related to empowerment of FDPs. Authors utilize a variety of qualitative methods including observations (following training sessions of social workers) and self-evaluations of the professionals. (Cambridge & Williams, 2004.)

The latter study, contrary to all other papers scrutinized here, can be considered as a 'meta-evaluation' while the author evaluates the evaluations of UNHCR refugee programs form an anthropological perspective. Empirically, the meta-evaluation was primarily based on anthropological fieldwork (participant observations and interviews) among refugees in Guinea. (Kaiser, 2004.) As a result of her fieldwork, the author makes a strong case for participatory methods in both planning and evaluating refugee programs: '[W]hen one restricts one's attention to the narrow practical parameters of humanitarian action, one neglects the wider, possibly causal, social, economic and political processes which have generated the need for them, and which function as their context […]. The extent to which it makes sense to do this is arguable, and may not, more importantly, reflect the priorities of affected populations. Defining a refugee operation as successful because it meets the expectations of a self-selected group of practitioners and evaluators, while neglecting to ask whether it provides a suitable response to the problems identified by refugee groups is tautologous and unhelpful.' (Ibid., pp. 202.) In other words, according to the author, participation of refugee populations and other hands-on stakeholders is necessary to conduct sustainable program evaluations.

3 Seven contexts of TAIS and their actor-oriented evaluation criteria

In this chapter, each of the seven RAISD contexts and TAIS are presented. The chapter is partly based on achievements of work packages four and five. In the following, information is provided on each context, their TAISs and actor-oriented criteria to evaluate them according to local stakeholders. As expected, local actor-oriented criteria were manifold. The stakeholders, mostly ARU members and interviewed FDPs in different contexts rarely provided explicit and detailed directions on how to evaluate certain practices. Instead, many of their accounts include ideas on what types of issues should be considered while implementing the pilots and what should be achieved with them. Thus, for the sake of clarity and to serve the purposes of integrated evaluation, actor-oriented criteria are divided into two categories: process and outcome evaluation criteria.

3.1 Spain

In Spain, dealing with forced displacement is highly dependent on various NGOs and individual citizens who are committed to human rights protection and welcoming new arrivals to the Spanish society. NGOs have adopted a strong role as a governmental subcontractor, while they are implementing (with scarce resources) public policies targeted to forcibly displaced persons. Within this system, distinct vulnerabilities are rarely recognised explicitly in the reception and inclusion system. Specific needs of FDPs are thus rarely met. FDPs exposure to hazards of discrimination, abuse or exploitation are rarely diminished. Inadequate services in reception system might further

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 822688.

produce vulnerabilities and expose FDPs to difficult situations. The living conditions FDPs face in the reception system are not fully acknowledged. The effects of dispersed families, low quality living conditions, lack of various resources, poor access to assistance and discrimination hiders the inclusion of many FDPs. The stakeholder data revealed several groups susceptible to vulnerabilities in the Spanish context such as single parent women, LGTBIQ-category, illiterates, trafficked people, unaccompanied minors, stateless people and so on. Consequently, in Spain, there were plenty of possibilities in which sub-group and vulnerability to focus on.

After a series of discussions and planning sessions with local stakeholders, the aim of Spanish TAIS is to promote the socioeconomic inclusion of Sub-Saharan female refugees or asylum-seekers by developing a series of trainings to support women's entrepreneurial skills. The objective is thus to promote participants knowledge and skills to create a cooperative or other kind of association that support their economic autonomy, and, possibly, that they make the first actions in the direction of entrepreneurship. The TAIS has a strong emphasis on social participation. Women are seen as protagonists; as key actors at the centre of the action.

To evaluate the TAIS, local stakeholders in Spain named several potential process and outcome criteria. The source of each criteria is written in brackets. Stakeholders refer to NGO representatives and other professional organizations while FDP refers to potential beneficiaries of the TAIS.

- Process criteria
    - Practices should be empowering people, not re-victimize them (FDP)
    - Hearing the voice of the included persons and avoiding paternalism (stakeholders)
    - Following and evaluating the process (stakeholders)
    - Individual, tailored services while the problem is lack of personnel/resources and domination of standardized services (stakeholders)
    - Human beings are the focus of attention (stakeholders)
    - In the case of vulnerable people among the FDP, inclusion must be a long procedure, there are not short-term processes, not financial standards (stakeholders)
    - Accessibility of the intervention (stakeholders)

- Outcome criteria
    - Aiming for inclusion (over integration), inclusion refers to rights, participation, empowerment and availability of information (stakeholders)
    - Equal access to health services, education and political participation (stakeholders)
    - Promoting the autonomy of individuals (stakeholders)
    - Co-operation with local stakeholders (stakeholders)
    - Promotion of diversity in the form of living together, not just to coexist (stakeholders)
    - Producing political activism & counteracting hate-speech (stakeholders)
    - Affecting the legal sphere since it is easier/faster to change than social/cultural spheres (stakeholders)

3.2 Italy

Over the past thirty years, Italy has transformed from a country of emigration that exported millions of Italians around the world, to a country of immigration. While this relatively new migratory flow began to modify the Italian ethnicities and cultures, the country was unprepared to deal with these global changes. Little progress has thus been made in terms of legislation and policies. Moreover, media coverage has depicted immigration as an assault over Italian coasts making immigrants the scapegoats of the many troubles of Italian society.

Currently, Italy functions as an entry region for many of those trying to make the risky journey from Northern Africa to Europe. So far, the Italian and European Union's strategy has been to take extreme measures to block the migration flows while probably violating several international agreements. Despite the strict border control mechanisms, Italian reception system is overcrowded and poorly resourced particularly for asylum seekers while better treatment is guaranteed for unaccompanied minors and those who are granted international protection. As for the reception of FDPs, only weak national legal and regulatory frameworks exist. Local practices thus tend to vary a lot. Concrete practices are often based on locally available resources, networks and attitudes of potential stakeholders. Moreover, the multiple and intersectional notion of vulnerability, the idea according to which an individual can be subjected to deprivations and discrimination based on several social categories simultaneously, is rarely recognised in the Italian society. Most of the interviewed FDPs in RAISD belong to several communities and this diversity, while being a source of strength, can expose a person to direct and indirect discrimination.

In this context, the Italian TAIS seeks to target female refugees and asylum seekers and to provide a novel platform for fostering their education. The TAIS is designed for 20 participants and aims to foster a bottom-up and participatory approach. Concretely, this means a shift from an induced trainings (top-down approach) to a selective approach in developing a trainings in which several optional learning modules are offered and participants can choose contents that are attractive for them. The education is provided by trainers and mentors sharing a similar background with potential beneficiaries.

To evaluate the Italian TAIS, the local stakeholders stressed the importance of the following process and outcome criteria:

- Process criteria
  - The practice enhances and recognizes previous learning experiences (stakeholders)
  - The practice follows a culturally sensitive and gender-oriented approach (stakeholders)
  - The practice is provided in accordance with institutional capacities (stakeholders)
  - The practice and procedures are easily comprehensible and their objectives are clearly defined (stakeholders)
  - The practice fosters comprehensive interdisciplinary of territorial Networking and Cooperation (stakeholders)
  - The practice promotes an increase in female participation (stakeholders)
  - The practice is society based for a more effective and sustainable response by the whole community (stakeholders)
  - The practice is organized in such a way as to ensure efficiency in management (stakeholders)
  - The practice design is based on operational flexibility (stakeholders)

- Outcome criteria
  - Practices should improve living conditions (FDP)
  - They should improve language and vocational skills (FDP)
  - Increase the knowledge on the employment possibilities (FDP)

Coordinator contact: Dr. Rubén Fuentes-Fernández | Universidad Complutense de Madrid | Calle del Profesor José García Santesmases, 9. Ciudad Universitaria 28040 MADRID, Spain.
t: +34/ 91 394 7548 | e: rfuentes@ucm.es | w: www.ucm.es , grasia.fdi.ucm.es

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 822688.

- The practice facilitates individuals' adjustment with the national integration systems' environment (stakeholders)
- The practice promotes own ethnic community engagement (stakeholders)
- The practice prepares beneficiaries for employment (stakeholders)
- The practice enables income generation for target beneficiaries (stakeholders)
- The practice ensures fair accessibility to local inclusion-services (stakeholders)
- The Practices promotes social interaction of both, local and foreign HVG participants (stakeholders)
- The practice promotes personal and social participation of HVG in order to be included in local communities (stakeholders)
- The practice provides that affective and psychological assistance is guaranteed (stakeholders)
- The practice pursues a policy for orientation, mentoring and tutoring of beneficiaries (stakeholders)

3.3 Hungary

After the collapse of the Soviet regime, the democratization and liberalization development of Hungary came to a halt after the 2008 economic crisis which was followed by a right-wing conservative turn in politics resulting in the establishment of a political and societal system that identifies itself as the System of National Cooperation, building an "illiberal democracy". For the past ten years politics and political communication in Hungary has become increasingly ethnocentric and nationalistic, framing immigration as one of the major threats to the integrity of the Hungarian people. As a result, the country has become increasingly hostile toward immigrants, especially asylum seekers. There has been a series of legal and political measures endorsed with a thorough political and public communication campaign that resulted in the dismantling of the refugee reception and integration systems.

Being located on the external borders of the European Union and the Schengen zone, in 2015 and 2016, Hungary was one of the key destinations for both regular and irregular migration. During these years, the Hungarian government decided to build a fence on the Hungarian-Serbian (and later on the Hungarian-Croatian) border to prevent unauthorized border crossings. Moreover, the government passed several legal amendments aiming to reduce migration to Hungary as well. For example, Hungary designated Serbia as a safe country, therefore apprehended migrants could be sent back to there. Furthermore, Hungary made expedited asylum determination possible, while guaranteed only limited procedural safeguards for applicants. Also, border crossings without travel documents was declared a criminal offence.

At the end of 2018, asylum seekers are de facto detained in the transit zones along the Serbian border during the status determination process. People with vulnerabilities, such as families, children and unaccompanied children above the age of 14 are no exception, only unaccompanied children under 14 are placed in a special children's home. All applications submitted by people entering from Serbia are considered inadmissible, appeal against the fast-track procedure is limited to 3 days. All migrants apprehended in Hungarian territory without a legal right to stay shall be "escorted" to the other side of the border fence along the Serbian border.

Coordinator contact: Dr. Rubén Fuentes-Fernández | Universidad Complutense de Madrid | Calle del Profesor José García Santesmases, 9. Ciudad Universitaria 28040 MADRID, Spain.
t: +34/ 91 394 7548 | e: rfuentes@ucm.es | w: www.ucm.es , grasia.fdi.ucm.es

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 822688.

These measures resulted in a situation where the number of asylum seekers and those who receive international protection has decreased to a few hundred per year. Refugee reception centres are virtually empty and community-based accommodation of asylum seekers is no longer possible. Those who receive some form of international protection usually leave Hungary for another European country soon after their status is granted, those who remain in Hungary face enormous hardships. Care and support are provided only by a handful of professional migrant-specific NGOs.

In addition to border and asylum politics, there are several weaknesses in institutional and inter-institutional level of recognising the needs of forcibly displaced people in general. It is so because funding for such institutions is low, the political context is hostile, and the service providing structures are in a difficult situation. Therefore, many times the root of the problem is that a client with complex vulnerabilities could not receive support at the right place in the right time. Many people fall out of the scope of supportive services, because complex problems cannot be tackled by one service provider alone. Also, in many cases the vulnerabilities or traumas are caused by the asylum system.

Consequently, the objective of the Hungarian TAIS is to enhance refugees' embeddedness in institutional structures. Social workers and service providers analyze individual cases in order to build a "Trajectory monitoring toolbox" of helpful and tailored interventions for vulnerable refugees. It contains evidence-based "vulnerability profiles" as well as the description of the trajectories that these typical cases should follow in the institutional field, in order to prevent the accumulation of vulnerabilities and to reduce the isolation of FDPs.

To assess the TAIS in Hungary, the local stakeholders stressed the importance of the following process and outcome criteria:

- Process criteria
    - Participants' initiatives should be channelled (stakeholders)
    - Community-based solutions should be explored (stakeholders)
    - Discrimination of all kinds should be avoided (stakeholders)
    - Inter-institutional and interdisciplinary cooperation should be fostered (stakeholders)
    - An overarching network of service providers should be reached (stakeholders)
- Outcome criteria
    - Social isolation of vulnerable individuals should be reduced (stakeholders)
    - Awareness of social workers and stakeholders should be raised (stakeholders)
    - Skills and knowledge of social workers and stakeholders should be expanded (stakeholders)

3.4 Finland

Finland is a Nordic country with a relatively large public sector, generous welfare benefits and citizen rights for all those with a residence permit. Thus, those without a residence permit, such as asylum seekers, are in a distinctively vulnerable position in terms of rights and access to services. The scarcely resourced reception system tailored for asylum seekers provides only some very fundamental services and is detached from most public services. Moreover, reception centres are often located in peripheral areas which highlights the confining nature of the system. Asylum seekers are isolated in many ways from the mainstream society – legislatively, institutionally, spatially and socially.

Coordinator contact: Dr. Rubén Fuentes-Fernández | Universidad Complutense de Madrid | Calle del Profesor José García Santesmases, 9. Ciudad Universitaria 28040 MADRID, Spain.
t: +34/ 91 394 7548 | e: rfuentes@ucm.es | w: www.ucm.es , grasia.fdi.ucm.es
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 822688.

Further, the influx of asylum seekers in 2015 and 2016 severely jammed the Finnish asylum processing system. Today, thousands of asylum seekers still wait their decision even five years after their arrival. The protracted uncertainty has been a heavy burden severely hampering the wellbeing of many people and adding another layer of isolation, a mental one. Moreover, an essential feature of the Finnish vulnerability context for asylum seekers emanates from the strict European border control. Only a selected group of people are able to make the expensive and dangerous journey through irregular routes to reach Nordic region. Consequently, the bulk of asylum seekers in Finland are young people, mostly single men.

Despite the confining nature of the Finnish reception system and lack of resources, some groups' vulnerabilities are recognized in the legislation and in institutional practices and they are provided relatively extensive services. These groups include unaccompanied minors, victims of human trafficking and asylum seekers suffering from more severe mental problems. Moreover, children in compulsory school age (6-17) are entitled to enrol to the education system. Thus, according to the stakeholder interviews and focus group discussions particularly two large groups of asylum seekers are left unrecognized: young asylum seeking single (and often lonely) men and (single) parents with small children.

In this context, two TAISs are developed. First, an online forum for Finnish-speaking volunteer men and asylum seeking men is tested. Men are invited to join the forum to discuss the experiences of asylum seekers and everyday life in the Finnish society. The discussions will be moderated and, when necessary, translated. Each pilot cycle will last approximately five weeks. The aim of the TAIS is to promote inter-group cohesion and foster the cultural learning and language acquisition of asylum seeking men. Quite often, asylum-seeking men lack the opportunities to use and learn everyday language with Finnish-speaking men. Particularly in smaller towns and rural areas there are few possibilities in contacting Finnish-speaking peers. Moreover, according to several asylum seeker interviews, the language tuition within the reception centres is not enough, takes place only in classrooms and is too repetitive due to constantly changing (and voluntary) teachers and learners and the huge diversity of educational backgrounds different asylum seekers have.

The aim of the second TAIS is to develop childcare services that already exist in some reception centres. A local team (ARU) will be established within two reception centre providing comparatively wide-ranging child-care services for small children. The aim is to observe and further develop the child-care services throughout the developmental cycles by producing knowledge both from the perspective of service providers and parents. Each cycle will last approximately six weeks. As a result, a model for child-care services for reception centres will be established. In most Finnish regions, asylum seeking families do not have the right to early childhood education in municipal day-care centres. Moreover, the level of child-care services provided varies significantly between different reception centres and this has repercussions on both parental and child wellbeing.

To evaluate the TAIS in Finland, the local stakeholders named the following process and outcome criteria:

- Process criteria
    - Listening to FDPs and integrating their ideas (stakeholders)
    - Listening the service providers and knowing what is suitable for them (stakeholders)
    - Adapting the pilot/intervention into local contexts, the conditions in different receptions centres may vary a lot (stakeholders)
    - Long processes, sustainability (stakeholders)

Coordinator contact: Dr. Rubén Fuentes-Fernández | Universidad Complutense de Madrid | Calle del Profesor José García Santesmases, 9. Ciudad Universitaria 28040 MADRID, Spain.
t: +34/ 91 394 7548 | e: rfuentes@ucm.es | w: www.ucm.es , grasia.fdi.ucm.es

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 822688.

- Targeted and tailored practices for different groups (stakeholders)
- The needs of different asylum seekers are recognized while accommodating them and gender, special needs/vulnerabilities, families and previous qualifications are recognized (stakeholders)
- Respecting the seven principles of The Red Cross: humanity, equality, impartiality, independence, voluntariness, unity and universalism (stakeholders)
- Co-operation with other reception centres and local actors are established to develop the quality of services (stakeholders)
- Work and study activities (in reception centres) are well organized and wide-ranging (stakeholders)
- Knowledge on customer satisfaction is produced and used to develop the work, several channels for giving feedback are used (stakeholders)

- Outcome criteria
  - A need to support families with children (stakeholders)
  - Asylum seekers receive orientation/introduction to the facilities and services (stakeholders)
  - The number of educated voluntary workers increases, centres are able to keep their old volunteers, asylum seekers are recruited as voluntary workers and voluntary work is wide-ranging (stakeholders)
  - Individual needs of asylum seekers are recognized and acknowledged in transfers to other reception centres (stakeholders)
  - Increase the autonomy of asylum seekers (FDP)
  - Decrease the confinement of asylum seekers into reception centre facilities (FDP)

3.5 Turkey

Turkey has a hybrid culture that combines both eastern and western cultures. Moreover, Turkey inhabits numerous ethnic and lingual groups. According to the latest available figures from the Turkish Directorate General of Migration Management (DGMM) there are currently over 4 million foreign nationals present in Turkish territory seeking international protection. Most are Syrians (3,667,435 individuals) who are granted temporary protection status, while according to UNHCR 368,230 asylum seekers and refugees from countries including Iraq, Afghanistan, the Islamic Republic of Iran and Somalia constitute another significant group of foreign nationals seeking Turkish humanitarian and legal protection. The Republic of Turkey is a party to the 1951 Refugee Convention and 1967 Protocol. However, Turkey is still maintaining the geographical limitation to the 1951 Convention while accepting refugees only from European territory. Moreover, Turkey strongly prefers temporary protection and resettlement in third countries as a durable solution over integration strategies. This means that there are huge numbers of forcibly displaced people in Turkey without recognized statuses, with limited rights and restricted access to various services. Despite this, FDPs in Turkey are provided free education and health care.

After the prolonged crises in several Middle-East countries, and despite the official policies, many FDPs in Turkey face a situation in which they seek ways to integrate themselves to local communities. Therefore, there is a need for learning the language, recognition of rights, access to services and finding places of study and employment opportunities. Currently, people suffer in many ways from their temporary statuses and lack of rights. The situation leads to uncertainty for future planning regarding accommodation, education and employment. Usually, FDPs need to resort to underpaid work, insufficient housing and poor services.

Coordinator contact: Dr. Rubén Fuentes-Fernández | Universidad Complutense de Madrid | Calle del Profesor José García Santesmases, 9. Ciudad Universitaria 28040 MADRID, Spain.
t: +34/ 91 394 7548 | e: rfuentes@ucm.es | w: www.ucm.es , grasia.fdi.ucm.es
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 822688.

As a consequence of the above structural conditions, the TAIS in Turkey aims to ensure that FDPs access and participation in basic daily life practices can be systematically followed through monitoring by host community social workers. Through monitoring, it will be possible to determine in which areas additional activities are required and also where necessary support is needed. Secondly, since forcibly displaced people cannot access and might be unaware of the existing possibilities, monitoring can translate in to an advocacy work as well.

To evaluate the TAIS in Turkey, the local stakeholders named the following process and outcome criteria:

- Process criteria
    - Diversity – respect of cultural, sexual, gender and religious heterogeneity (stakeholders)
    - Anti-discrimination – prevention of hate speech and avoidance of racial exclusive terminology (stakeholders)
- Outcome criteria
    - Social cohesion in informal spheres – mutual respect and ability to live together (stakeholders)
    - Integration in formal spheres – from guests to residents (stakeholders)
    - Language acquisition – learning Turkish (stakeholders)

3.6 Lebanon

Lebanon is not a party to the Convention relating to the Status of Refugees of 1951 or its 1967 Protocol. It has not adopted any domestic legislation specifically addressing the status of refugees. Refugee status is at present determined mainly by the provisions of a Memorandum of Understanding (MOU) signed between Lebanon and the UNHCR. The contract provides a mechanism for the temporary residence permits to asylum seekers. Therefore, the legal status of refugees in Lebanon lacks certainty. The existing legal instruments dealing with this issue have been criticized as inadequate and insufficient. A 2010 report by the UNHCR states that "[r]efugees enjoy few, if any, legal rights in Lebanon." However, in formal level, Lebanon's domestic legislation allows one to apply asylum and fosters the principle or non-refoulement.

Lebanon hosts a large number of refugees. In a report issued by the Humanitarian Aid and Civil Protection department of the European Commission in October 2015, it was estimated that there were about 1.1 million Syrians, 295,000 Palestinians, and 17,000 Iraqis in Lebanon, while UNHCR estimated the number of Syrian refugees alone to be 1,835,840 in 2015.

During the RAISD-project, interviewed FDPs in Lebanon have been refugees from Syria. None of the interviewees had enjoyed the help from public bodies or NGOs. They had resorted to other people and tried to exploit the precarious work opportunities available. Women-headed households and households with pregnant women were among the ones in most vulnerable position. A high proportion of females reported the prevalence of women-headed households (widows), while many reported the presence of abuse and exploitation in their communities. Moreover, interviewees reported about various barriers for inclusion in the hosting communities as well as about the high prevalence of refugee children not attending schools.

As for medical needs and health care services, several interviewees reported about unmet needs. Due the economic situation in Lebanon, interviewees identified with serious illness, mental disorders, and other serious forms of psychological, physical, or sexual violence. All interviewees reported the lack of any health services. Due

Coordinator contact: Dr. Rubén Fuentes-Fernández | Universidad Complutense de Madrid | Calle del Profesor José García Santesmases, 9. Ciudad Universitaria 28040 MADRID, Spain.
t: +34/ 91 394 7548 | e: rfuentes@ucm.es | w: www.ucm.es , grasia.fdi.ucm.es
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 822688.

to above adversities, interviewees rarely saw their future in Lebanon. Instead, many of them had dreams of migrating further to 'Western' countries.

In the above context, the TAIS developed in Lebanon addresses immediate social and health related issues. The program aims to have a positive impact on the Syrian people in camps while delivering a holistic Corona Virus awareness camping to Children, pregnant women, elderly people, malnourished people, and people who are ill or immune-compromised. LIU Health committee will work to enrich all stakeholders with the coping tools needed to enhance them in developing their awareness and minimize the challenges in facing the COVID 19 pandemic trauma.

To evaluate the TAIS in Turkey, the local stakeholders named the criteria below. All of them can be considered as outcome criteria focusing on the results of the developmental work.

- Outcome criteria
    - Empowerment of FDPs living in refugee camps
    - Enhancement of their capacity building
    - Development of their language skills
    - Stabilization of their emotional wellbeing

3.7 Jordan

Jordan is a stable and relatively wealthy country on the East Bank of the Jordan River. As a neighbouring country of many conflict-ridden areas, Jordan has an extensive history of hosting large refugee populations. After the 1967 war with Israel and Iraq's 1990 invasion of Kuwait, there were sudden and massive influxes of Palestinian Arab refugees, who now make up more than two-thirds of the population. In 1996, 1,359,000 Palestinian refugees living in Jordan were registered with United Nations; 250,000 Palestinians continue to live in ten refugee camps. Today, Jordan is among the countries affected most by the Syrian crisis, with the world's second highest share of refugees compared to its population. Jordan has 89 refugees per 1,000 inhabitants [UNHCR, 2018]. Most of Syrian refugees are in the northern governorates of the country: Amman, Irbid and Mafraq. Recently, Sudanese and Yemeni refugee populations have also started to grow, as has the presence of migrant workers from Egypt, and others from African and Southeast Asian countries.

Although Jordan has not signed the 1951 Refugee Convention, it has recognized the refugee status granted by UNHCR. UNHCR refugee status grants holders the ability to stay in Jordan for the period of six months and gives them the ability to work through the issuance of work permits. Moreover, the UNRWA mandate with the Jordanian government has granted the Palestinian refugee's five-year passports. Based on the Memorandum of Understanding (MOU) signed between UNHCR and Jordan in 1998, refugees can seek asylum in Jordan up to six months, which can be extended by refugees. This understanding has resulted in the open-door policy currently adopted by the Jordanian government towards both Syrian and Iraqi refugees. Thus, Jordan's government considers refugees as guests in the country.

In the formal level, several rights of some FDP groups are recognized. For instance, family reunification is granted in certain situations and Syrian refugees are granted access to public services like schools. However, refugee households are still live below the poverty line. Extreme poverty causes malnutrition, health related problems, homelessness and inability to enrol into educational institutes. Consequently, a huge number of FDPs in Jordan

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 822688.

live in highly marginal positions. In the RAISD interviews it was found that there are five essential challenges and unmet needs: extreme poverty, lack of educational opportunities, malnutrition, mental and health related problems and various daily needs. Consequently, Jordan was perceived as a transit country by many of the interviewees.

In Jordan, to alleviate the above problems, a Psychosocial Refugee Support Forum (PRSF) will be developed. PRSF aims to promote the well-being of individuals and prevent the need for medical support through non-clinical interventions. PRSF opens the door for the participation of stakeholders working with the Jordanian research team along with sharing experiences, exchanging data, and building up training programs to support the target group. Five different topics will be implemented by the PRSF through supporting: Economic Empowerment and enhancing financial knowledge, social support and refugee status, health and legal awareness, and empowerment and psychological well-being. Moreover, RAISD-YU team attempts to create Refugee Support System portal "RSS", which was established to spread awareness of the benefits of psychosocial support, provide the support for refugees, and collect and preserve data.

To evaluate the TAIS in Jordan, local stakeholders named the following process and outcome criteria:

- Process criteria
    - Ensuring that support is targeted to vulnerable groups
- Outcome criteria
    - The promotion of registration of FDPs (legal status)
    - Providing psychological support
    - Raising awareness on the rights of FDPs
    - Assisting and supporting FDPs
    - Encouraging organizations to provide cash assistance to FDPs
    - Promoting effective and high-level health services
    - Supporting in costs of health services
    - Preventing substance abuse
    - Promote students' integration in their schools
    - Promoting young FDPs admission to tertiary education (scholarships, assistance in costs)
    - Promoting the ability of tertiary education institutes to deliver quality teaching to FDPs (e.g. monitoring and mentoring)
    - Facilitating work permissions for FDPs

4 Integrated evaluation criteria

4.1 Basic principles – a return to the realistic tradition of evaluation research

Realist evaluation is a theory-driven enquiry. However, instead of 'grand narratives' or macro-level theories of sociology, the aim is often to conceptualise the mechanisms through which a program/intervention enables people to make changes happen. Thus, a realist evaluator is sensitive to contexts and their distinct features while aiming to build knowledge with multi-method strategies and by engaging with the field of study and stakeholders. Thus, for a realist evaluator, 'context is not unwelcome noise, not a confounding variable to be controlled for – but an integral part of a programme' (Pawson, 2013, pp. 36). In contrast to experimental settings dominating the

Coordinator contact: Dr. Rubén Fuentes-Fernández | Universidad Complutense de Madrid | Calle del Profesor José García Santesmases, 9. Ciudad Universitaria 28040 MADRID, Spain.
t: +34/ 91 394 7548 | e: rfuentes@ucm.es | w: www.ucm.es , grasia.fdi.ucm.es

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 822688.

game of evaluation studies, assessing an intervention does not need to exclude the effect of all unknown factors by randomisation, but rather take them as an object of investigation. For a realist evaluator, it is essential to answer also why questions instead of merely scrutinizing whether the program works or not (cf. Patton, 2014; Kivijärvi et al., 2019). Consequently, realist evaluation is congruent with the principles of RRI: both highlight the importance of stakeholder and beneficiary perspectives instead of embracing universalist and top-down approaches.

At this point it is probably evident that in realist evaluation contexts are important. Pawson (2013, pp. 37) defines four contextual layers that need to be taken into account when assessing a program. The first layer is an individual one referring to characteristics, capacities and motivations of various stakeholders and end-users. In RAISD, it is of high importance to engage with the actual people who provide the intervention and those who could benefit from it. The second layer is interpersonal and refers to individual level relationships and their quality. In RAISD, the relationships between, for instance, professional service providers and FDPs probably affect the outcomes of interventions. The third layer is an institutional one. Interventions are always implemented in the context of some institution whether it is a school, hospital, an NGO or a reception centre. An evaluator must be familiar with the institution, its regulation and limits. The fourth layer is infrastructural as it refers to macro-level societal conditions, legislation, dominating discourses, institutional hierarchies and cultural structures.

In other words, being a realist evaluator requires a lot. One should master both hands-on micro-level information in a specific setting and more macro-level knowledge on societal conditions and institutional preconditions for a certain program. This is an objective of the RAISD consortium and is targeted by specific work packages. The focus of WP4 is on infrastructural and institutional layers in describing the current policies, institutional actors and programs in each context. WP3 holds a key position in providing knowledge on individual and interpersonal layers by interviewing various stakeholders, potential service providers and FDPs in different positions. All this contextual knowledge helps RAISD evaluators in answering the key questions of realist evaluation: what might work for whom in which contexts?

4.2 Integrated evaluation criteria in RAISD

In terms of contextual layers presented below, RAISD pilots will primarily take place in individual and interpersonal layers in contrast to institutional and infrastructural ones. Pilots will thus be rather small scale, involve people working in the field and they will be implemented in real life settings. To quote deliverable the 5.2 on Tailored Attention and Inclusion Strategies:

> Macro level is the context to keep in mind all the time and recommendations can be aimed at the macro level but TAIS should be implemented on the meso/micro level. If there is an obstacle found on the macro level, which makes a given TAIS proposal unfeasible, it should be modified to something feasible.

In practice, interventions at the institutional level could target an NGO or a bureau and aim to improve its accessibility (e.g. development of information flow or more suitable opening hours), procedures (e.g. new methods of service delivery), skills/attitudes/knowledge of staff or relations with other organisations. Individual and interpersonal levels can include a wide variety of interventions. They might aim in awareness raising, improving relations between service providers and FDPs, strengthening intercultural skills, offer peer support or adjust services to better meet FDPs needs.

Coordinator contact: Dr. Rubén Fuentes-Fernández | Universidad Complutense de Madrid | Calle del Profesor José García Santesmases, 9. Ciudad Universitaria 28040 MADRID, Spain.
t: +34/ 91 394 7548 | e: rfuentes@ucm.es | w: www.ucm.es , grasia.fdi.ucm.es
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 822688.

In the following, a more detailed account on RAISD integrated evaluation criteria will be provided. First, a general framework is set including the key issues to keep in mind while implementing the TAISs. Second, a more detailed description of the integrated criteria is delivered to cover evaluation activities during the different phases of implementation. These criteria are derived from actor-oriented criteria presented in the chapter 4.

4.2.1 Integrated evaluation criteria: A general framework

The structure of the piloted activities is organised around the ARUs iterative work. The implementation of the interventions will have a cyclic character, gradually refining the practices in each vulnerability context to ensure their adaptation and productivity. In practice, RAISD pilots will be implemented in three cycles. During the cycles, they will be developed further according to stakeholder and end-user feedback. Throughout the cycles, the focus is on increasing the pilot's feasibility for users.

The first and probably the most important evaluation criterion concerns, not surprisingly, vulnerability contexts. The key to planning successful interventions is that they fit the contexts in which they are implemented in. In other words, apt definitions of vulnerability contexts ensure that suitable questions are asked, promising pilots are set up, relevant vulnerabilities are targeted and that FDP needs are met – contextual knowledge provides the rationale for each TAIS that is be implemented during the project.

The second general evaluation criterion is related to mechanisms. The task is to explore the mechanisms through which interventions promote desired changes. The key here is to investigate interventions as processes. To do this, an evaluator needs to be engaged with the people involved in implementing the TAIS. Very often this requires compiling qualitative data in the form of observations and interviews. Concrete questions need to be posed: How is pilot perceived among different stakeholders? What happens during the pilots? How could they be improved? How much time and effort it requires from stakeholders? Process evaluation should be given the priority during the first cycles in order to ensure that pilots are implemented in feasible and ethically sustainable ways and that they provide useful resources for stakeholders.

In line with the RRI principles, it is important to include different stakeholders while carrying out the process evaluation. This imperative includes the recognition of the diversity of potential perspectives. Depending on the TAIS, the diversity might stem from differences in gender, ethnicity, ability or from any other social position. Moreover, service providers and end-users might have divergent views on how the TAIS should be implemented.

The third general evaluation criterion is related to outcomes. Towards the end of pilot cycles, and after the process of delivering the TAIS has been refined during previous cycles, it is necessary to determine the possible harms/benefits the intervention provides for stakeholders. At this stage, longitudinal datasets are preferred in order to estimate the change interventions might have provoked. In practice, this often means follow-up surveys or retrospective interviews among potential beneficiaries of the interventions. Further, outcome evaluation can include a cost estimation as well. This would include information on working hours and materials needed and relating them to the benefits gained.

Outcome evaluation should be given the priority during the third and final pilot cycle. As with process evaluation, also in outcome evaluation it is of essential importance to deploy the RRI principles. Outcomes might not be the same for all. Differences might be observed in outcomes between service providers and end-users and between different sub-groups among them. Again, these differences might stem from varying societal statuses or positions

Coordinator contact: Dr. Rubén Fuentes-Fernández | Universidad Complutense de Madrid | Calle del Profesor José García Santesmases, 9. Ciudad Universitaria 28040 MADRID, Spain.
t: +34/ 91 394 7548 | e: rfuentes@ucm.es | w: www.ucm.es , grasia.fdi.ucm.es

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 822688.

within an institution that the pilot is targeting. In outcome evaluation, the potential diversity of perspectives needs to be recognised. In realist evaluations, the aim is to study for whom the program works and for whom it does not.

Table 2. General evaluation framework

General criteria	Tasks	Methods	Timing
Defining needs	Conceptualising vulnerabilities and vulnerability contexts	Literature, statistics and stakeholder and FDP interviews	Before starting the pilots
Defining mechanisms	Determining what happens during the pilots and how they are implemented and perceived	Stress on qualitative data: interviews, participatory observation, action research	Mainly during the first two pilot cycles
Defining outcomes	Determining the level of change the piloted activities provoke and estimating their costs	Stress on quantitative data: mostly follow-up surveys but also retrospective interviews	Mainly before and after the last pilot cycle

The three general evaluation criteria defined above are summarised in table 2. This general framework is, to a great extent, in line with the so-called SMART criteria set for 'effective TAIS' in the deliverable 5.2. According to the latter criteria pilots should be Specific, Measurable, Attainable, Relevant and Time-based. The criteria on specificity and relevancy correspond to needs definition: piloted activities should be clearly targeted and related to observed vulnerabilities. Further, attainability and temporality are related to mechanisms definition. Pilots should be viable and conceivable with the available resources and timeframes. Finally, measurability corresponds to criteria on outcome definition while the change should be measured by using relevant (quantitative) indicators.

4.2.2 From actor-oriented to integrated criteria

Integrated criteria are based on a careful analysis of actor-oriented criteria and set in the frames of general evaluation framework presented above. This sub-chapter serves a two-fold purpose. First, an analysis and selection of actor-oriented criteria will be presented. Second, the most relevant actor-oriented criteria will be adjusted to fit the integrated evaluation scheme of RAISD.

Determining integrated evaluation criteria

All country-specific actor-oriented criteria are listed in chapter 3. As the first step of the analysis, several alike criteria were combined in order to reduce their number and to formulate a more comprehensible totality. The second step was to rank both process and outcome criteria based on their frequency. Some criteria were mentioned by stakeholders in almost all partner countries while some were mentioned only once. As for process evaluation, the most frequently mentioned criteria were respecting FDP diversity, listening to FDPs and listening to service providers. As for outcome criteria, particularly wellbeing and inclusion of FDPs were frequently mentioned. In more detail, these included criteria such as increasing the autonomy and empowerment of FDPs and promoting their access to various institutions.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 822688.

After counting the frequencies, as the third step of the analysis, actor-oriented criteria were again ranked according to their viability in the framework of RAISD-project. Certain rarely mentioned criteria were excluded due to several reasons: some criteria could be merged (e.g. process criterion on anti-discrimination was included in criterion of respecting FDP diversity and outcome criteria on fostering FDPs autonomy and empowerment were fused), some were not targeted by any TAIS (e.g. outcome criteria on preventing substance abuse and fostering engagement in one's 'own' community) and some were difficult to assess or exceeded the scope and timeframes of the project (e.g. reducing hate-speech and changing the legislation). A detailed list on included and excluded actor-oriented criteria can be found in Annex 1.

As a result of the analysis, a set of both process and outcome criteria were established. These criteria are presented in Table 3. Obviously, the list is quite comprehensive and not all of them need to be used in evaluating every TAIS. There are differences between the aims of TAIS and how they will be implemented. These quite fundamental differences need to be considered while deciding the final evaluation criteria for each TAIS. Most pilots focus directly on promoting the wellbeing of FDPs while some target professional practices of service providers. Moreover, some pilots will be implemented in iterative fashion while others will follow step-by-step procedures. The former refers to a strategy in which the TAIS is implemented already during the first pilot cycle and adjusted according to evaluation results in the course of second and third cycles. In the latter strategy, the first and possibly even the second cycle is dedicated to planning the TAIS by continuing knowledge production and training relevant stakeholders while the actual implementation takes place only during the last cycle. Consequently, in the end, all partners follow the common evaluation frames but at the same time they need to adapt it according to their local needs.

Coordinator contact: Dr. Rubén Fuentes-Fernández | Universidad Complutense de Madrid | Calle del Profesor José García Santesmases, 9. Ciudad Universitaria 28040 MADRID, Spain.
t: +34/ 91 394 7548 | e: rfuentes@ucm.es | w: www.ucm.es , grasia.fdi.ucm.es

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 822688.

Table 3. Actor-oriented criteria to be included in the integrated evaluation criteria

Process criterion	Frequency
Respecting FDP diversity	16
Listening to FDPs	5
Listening to service providers	5
Fostering cross-professionalism	3
Flexibility and sensitivity to local context	3
Wide reach	2
Outcome criterion	**Frequency**
Increased autonomy	12
Increased professional skills / job market awareness / employment rate	9
Increased access to services/support/education	8
Improved living conditions	5
Promoted intergroup cohesion	5
Improved language skills	3
Increased knowledge on rights/available services	2
Increased cross-professional co-operation	1

An elaboration of integrated evaluation criteria

In the following, an elaborated framework for integrated evaluation criteria is presented. This means embedding and adjusting the selected actor-oriented criteria (table 3) to the general evaluation framework proposed in table 2. In the integrated evaluation framework, the actor-oriented criteria are not repeated in their 'raw' form. In order for the criteria to fit the partners' interests and the ideas of realist evaluation, they have been on one hand generalised into a more abstract set of categories, and on the other hand, elaborated further to determine the methods and datasets to be used in the process and outcome evaluation. A detailed list of actor-oriented criteria and their analysis can be found from Annex 1.

As for the **process evaluation**, when looking at actor-oriented criteria, both FDP and service provider perspectives are strongly emphasised. Criteria on respecting FDP diversity, listening to them and wide reach are directly related to potential end-users of the pilots. In evaluation practice, this set of criteria can be translated into categories of accessibility, acceptability and empowerment. In addition to FDP or end-user perspective, the standpoints of service providers are covered in the criteria on listening them, fostering cross-professionalism and pilots' flexibility and sensitivity to local contexts. In the integrated evaluation criteria, the notion of viability summarises the service provider perspective. Below, all four process evaluation categories are clarified further.

Coordinator contact: Dr. Rubén Fuentes-Fernández | Universidad Complutense de Madrid | Calle del Profesor José García Santesmases, 9. Ciudad Universitaria 28040 MADRID, Spain.
t: +34/ 91 394 7548 | e: rfuentes@ucm.es | w: www.ucm.es , grasia.fdi.ucm.es
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 822688.

- *Accessibility* means that pilots can be reached and used by people with various abilities and resources. Accessibility can be evaluated by compiling statistics on the profiles or social statuses of those who are reached and interviewing end-users on the ease of access, ease of use and non-discriminatory nature of the pilot implementation. Moreover, if possible, also potential end-users who are not reached can be interviewed. They are likely the key informants in providing the most critical information on accessibility.
- *Acceptability* refers to perceptions of the end-users and whether and to what extent the pilot is adopted by them. In the evaluation literature acceptability most often means determining how certain piloted activity is received by the beneficiaries (e.g. Sekhon et al. 2017). Even though the starting point in evaluating the acceptability is qualitative and subjective (how-question), it can be studied with quantitative and objective methods as well. Therefore, in addition to qualitative interviews in which the reception of the pilot is probed among end-users, an evaluator can use behavioural indicators on withdrawal and adherence rates (see e.g. Knowles et al. 2017).
- *Empowerment* takes the process evaluation one step further and comes closer to exploring mechanisms. Evaluating the empowering potential of pilots means to explore the resources or recognition they might provide for stakeholders in accomplishing the changes they aspire for. Provided resources might help people to change their behaviour in opening up new possibilities for them while being recognised might turn into changed patterns of thinking or relating to one's surroundings. In practice, assessing the empowering aspects of pilots requires subjective and qualitative data, most often in the form of interviews. An evaluator needs to know how pilots are perceived by the beneficiaries, how they are used and for what purposes.
- *Viability* refers not to an end-user but to a service provider perspective. Pilots need to be seen as realisable and appealing by service providers. If this is the case, it increases the likelihood that pilots are promoted in a way that end-users can benefit from them. Most importantly, viability is dependent on institutional and professional standards and ideologies and resources available. In other words, pilots should not be too burdensome for the practitioners and compatible with their professional orientations (see Kivijärvi et al. 2019). A further aspect in assessing viability is pilots' flexibility; whether practitioners are able to deliver them in ways that seems reasonable for them. Successful and relatively effortless cross-professional co-operation might be seen as a process toward more comprehensive and effective service provision as well. Viability of the pilots can be assessed with several types of datasets. It might be done by observing the frequency or intensity with which the pilot is used or provided. Moreover, qualitative interviews of service providers should be used as well.

A quite detailed set of different outcome criteria (table 3) also need be abstracted into a more comprehensive categories of outcome evaluation. As with process criteria, in the included actor-oriented outcome criteria both FDP and service provider perspectives are highlighted. However, the great majority of the actor-oriented outcome criteria focus on FDPs. These include increasing FDPs autonomy, improving their labour market position, securing their access to various institutions, improving their living conditions, promoting intergroup solidarity, improving language skills and knowledge on their rights in the host society. Below, these are transformed into three overarching categories: capability, inclusion and capacity. Service-providers have a quite marginal status in actor-oriented outcome criteria while only promoting cross-professionalism is included. However, service providers

Coordinator contact: Dr. Rubén Fuentes-Fernández | Universidad Complutense de Madrid | Calle del Profesor José García Santesmases, 9. Ciudad Universitaria 28040 MADRID, Spain.
t: +34/ 91 394 7548 | e: rfuentes@ucm.es | w: www.ucm.es , grasia.fdi.ucm.es

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 822688.

might be affected in other ways as well. This could mean increased skills in working with FDPs and promoting a more comprehensive understanding on forced migration and vulnerabilities. Consequently, competence will function as the overarching outcome category for service providers or practitioners.

- *Capability* is a vastly debated notion in the field of social sciences and one approach to wellbeing of people. In short, capability refers to actual possibilities people have to achieve the goals they aspire. This includes both the resources that people have and the circumstances in which these resources can be used. Thus, wellbeing is considered not only in terms of assets but as abilities to use and benefit from them in certain contexts. More concretely, this might refer to the kind of social life, level of health and lifestyle people wish to pursue. (E.g. Sen 1992; 1993). In this sense, the notion of capability relates to the notion of autonomy – the ability of individuals to pursue the goals that are the most desirable for them. Changes in capabilities can be traced with both qualitative interviews and by using established survey scales.
- *Inclusion* is a multifaceted concept. A narrow definition of the term refers to integration of certain minorities to mainstream institutions (such as disabled children attending regular classes). A broader definition denotes minorities' societal participation in general: social connectedness, labour market participation, educational degrees, community engagement and access to services. The actor-oriented criteria clearly relate to the broader definition of the notion. This type of inclusion can be assessed as a subjective and an objective outcome. Subjective inclusion relates to feelings of belonging to local communities and host society while objective inclusion can be measured as ties with majority representatives, service use and as an access to educational institutes and labour market. Due to the all-encompassing nature of the notion, each partner need to determine the most relevant aspects to be evaluated and targeted by their pilots. Some pilots might explicitly aim to increase intergroup cohesion while some others focus mainly on promoting educational opportunities.
- *Capacity* is a category related both to wellbeing and inclusion while referring to individual FDPs and their abilities to steer their lives and to participate socially and politically. In the actor-oriented criteria, capacity refers to knowledge and skills. The latter are professional and language skills while the former denotes to knowledge on individual rights and available services. In the outcome evaluation, capacity could also refer to a more in-depth understanding of the cultures and institutions of the host societies. Change in the capacities can be assessed with various methods. In order to acquire objective data on the potential progress of FDPs, language skills and know-hows can be tested. However, since tests are designed to produce top-down information they might be disempowering for many FDPs. Therefore, also subjective information should be composed by focusing on the perceptions of FDPs. This would mean surveying and interviewing pilot participants on how their skills and awareness have developed.
- *Competence* is related to institutions and service providers. The actor-oriented-criteria include a quite narrow view of competences since only cross-professional co-operation is mentioned. Cross-professionalism is more a means to achieve outcomes, not an outcome in itself. Consequently, it can be a target of process evaluation but not outcome assessment. Cross-professional work as a process may foster learning while practitioners from different backgrounds brainstorm together and need to set common goals and agree on the means to achieve them. Moreover, pilots might promote learning and

Coordinator contact: Dr. Rubén Fuentes-Fernández | Universidad Complutense de Madrid | Calle del Profesor José García Santesmases, 9. Ciudad Universitaria 28040 MADRID, Spain.
t: +34/ 91 394 7548 | e: rfuentes@ucm.es | w: www.ucm.es , grasia.fdi.ucm.es
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 822688.

knowledge accumulation also by introducing new ways of reflecting the service provision, encountering FDPs and providing first-hand knowledge on their living conditions. Transformations related to learning and possible changes in professional orientations are complex and quite unpredictable phenomena and thus not easily operationalised as survey items. Additionally, the number of service providers included in each pilot might be relatively modest. Therefore, while evaluating changes in competence, resorting to qualitative and retrospective interviews is probably a must.

Finally, in the table 4, a framework for integrated evaluation criteria is set including the four target categories for both process and outcome evaluation. Integrated evaluation criteria are congruent with the general evaluation framework presented in the table 2. After a careful analysis of vulnerability contexts (defining needs), chosen TAISs are implemented in each partner country. While implementing the pilots, particularly during the first and second pilot cycles, the focus will be on process evaluation and determining the mechanisms potentially generating the sought-after outcomes. Consequently, accessibility, acceptability, empowering aspects and viability of the TAISs are assessed. This requires resorting mainly to qualitative methods in the forms of interviews and requested feedback but also statistics on the number and profiles of end-users, the frequency of using the pilot and on the adherence and withdrawal rates. Process evaluation and exploring the potential mechanisms provide data on how to refine and develop the pilot further, and how to explain the possible outcomes.

In the last and third cycle of pilot implementation, a plenty of evidence-based developmental work has already taken place. Consequently, in the third cycle, pilots will be implemented in its refined and final form. At this stage, the main task is to evaluate the outcomes in the forms of changes in capabilities, inclusion, capacities and competences. Since observing the changes is crucial when evaluating the outcomes, there is a need of longitudinal assessments. This means conducting baseline and follow-up surveys or retrospective interviews. Overall, the focus of outcome evaluation is on FDPs and changes in their capabilities, their social participation and in their abilities to cope in the fields of the host societies. An important aspect of outcome evaluation is also assessing the impact of TAIS implementation on service providers' knowledge and abilities. Moreover, a calculation of costs caused by the TAIS implementation need to be conducted (not included in the table). This requires compiling information on the human resources (e.g. working hours) and material expenditures (e.g. leaflets, applications and premises) in order to provide rough estimates on the 'cost-effectiveness' of the pilots.

Coordinator contact: Dr. Rubén Fuentes-Fernández | Universidad Complutense de Madrid | Calle del Profesor José García Santesmases, 9. Ciudad Universitaria 28040 MADRID, Spain.
t: +34/ 91 394 7548 | e: rfuentes@ucm.es | w: www.ucm.es , grasia.fdi.ucm.es

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 822688.

Table 4. Framework for Integrated evaluation criteria

PROCESS EVALUATION	POSSIBLE DATA & METHODS
Accessibility The level of ease TAIS is reached by beneficiaries	• Statistics: Number of beneficiaries & distributions of them (according to gender, ethnicity, ability etc.) • Qualitative feedback from beneficiaries: Ease of access/use & non-discrimination • Qualitative feedback from stakeholders: Potential beneficiaries not reached (and why it was so)
Acceptability How TAIS is perceived by the beneficiaries?	• Statistics: Adherence (e.g. intensity of use or level of attendance) and withdrawal rates (if possible, according to gender, ethnicity, ability etc.) • Qualitative feedback from beneficiaries or stakeholders: How TAIS match needs? • Qualitative feedback from beneficiaries or stakeholders & observations: How TAIS is received?
Empowerment What types of resources TAIS provides or what types of changes it provokes for the beneficiaries?	• Qualitative feedback & observations: How pilots are used and for what purposes? o Might be different than originally planned • Qualitative feedback & observations: What kind of resources and/or recognition TAIS provides for the beneficiaries? o Might be different than originally planned
Viability Feasibility of the TAIS from the service provider perspective	• Statistics: Frequency of using/providing TAIS • Statistics: Increased cross-professional contacts as a result of TAIS • Qualitative feedback: Ease of providing TAIS • Qualitative feedback: How TAIS match with professional/institutional culture?
OUTCOME EVALUATION	POSSIBLE DATA & METHODS
Capability To what extent TAIS influences beneficiaries' perceived autonomy and coping in the context of host communities?	• Follow-up surveys and/or retrospective interviews: Change in the level of autonomy (e.g. perceived ability to cope in one's surroundings and belief in one's abilities in the future) • Follow-up surveys and/or retrospective interviews: Change in the level of perceived coping skills (e.g. communicating and awareness on one's possibilities)
Inclusion To what extent TAIS influences societal participation of beneficiaries?	• Follow-up surveys and/or retrospective interviews: Change in social networks & community engagement • Follow-up surveys and/or retrospective interviews: Change in employment or educational status and change in the level of access to services and institutions
Capacity To what extent TAIS influences competences/skills of beneficiaries?	• Follow-up surveys and/or retrospective interviews: Knowledge on work and educational opportunities, and on the rights and available services • Follow-up surveys and/or retrospective interviews and/or tests: Change in language skills and societal/cultural knowledge
Competence To what extent TAIS influences competences/skills of service providers?	• Retrospective interviews/feedback: Learning instigated by new methods of service provision or cross-professional co-operation • Retrospective interviews/feedback: Professional/institutional transformations as a result of providing the TAIS

The integrated evaluation criteria provide a framework for assessing each TAIS implemented in the context of the RAISD-project. Even though the foundation of the integrated evaluation is criteria defined in highly diverging surroundings, actor-oriented criteria were surprisingly similar in highlighting same types of aspects such as

Coordinator contact: Dr. Rubén Fuentes-Fernández | Universidad Complutense de Madrid | Calle del Profesor José García Santesmases, 9. Ciudad Universitaria 28040 MADRID, Spain.
t: +34/ 91 394 7548 | e: rfuentes@ucm.es | w: www.ucm.es , grasia.fdi.ucm.es
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 822688.

respecting the diversity and promoting the autonomy of FDPs. Despite the similarities in evaluation criteria, the concrete TAIS piloted in each partner country will differ in many respects. For individual evaluators, it is thus not necessary or maybe even reasonable to punctually follow the integrated criteria and its methods to a full extent. Moreover, individual partners should add their own evaluation criteria outside the integrated evaluation framework when necessary. In the end, the notion of framework describes the integrated evaluation criteria aptly: It provides a foundation for all partners which should be adjusted to local needs.

Coordinator contact: Dr. Rubén Fuentes-Fernández | Universidad Complutense de Madrid | Calle del Profesor José García Santesmases, 9. Ciudad Universitaria 28040 MADRID, Spain.
t: +34/ 91 394 7548 | e: rfuentes@ucm.es | w: www.ucm.es , grasia.fdi.ucm.es

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 822688.

5 Bibliography

- Astbury, B., & Leeuw, F. (2010). Unpacking Black Boxes: Mechanisms and Theory Building in Evaluation. *American Journal of Evaluation*, 31(3), 363-381.
- Bauböck, R., & Helbling, M. (Eds.) (2011). Which indicators are most useful for comparing citizenship policies? RSCAS Working Paper 2011/54. European University Institute, Robert Schuman Centre for Advanced Studies, EUDO Citizenship Observatory.
- Beine, M. et al. (2016). Comparing immigration policies: An overview from the IMPALA database. *International Migration Review*, 50(4), 827-863.
- Bhaskar, R. (1986) *Scientific Realism and Human Emancipation*. New York: Verso.
- Busch, N., & Fong, R., & Cook H.L., & Faulkner, M., & Mahapatra, N. (2007). An evaluation of the central Texas coalition against human trafficking. Austin: The University of Texas. Retrieved from http://www.utexas.edu/ssw/dl/files/cswr/institutes/idvsa/publications/evaluation_of_trafficking-2007.pdf
- Caliber (2007). Evaluation of comprehensive services for victims of human trafficking: Key findings and lessons learned. Fairfax, VA: National Criminal Justice Reference Service. Available at: https://www.ncjrs.gov/pdffiles1/nij/grants/218777.pdf
- Cambridge, P., & Williams. L. (2004). Approaches to Advocacy for Refugees and Asylum Seekers: a Development Case Study for a Local Support and Advice Service. *Journal of Refugee Studies*, 17(1), 97–113. https://doi.org/10.1093/jrs/17.1.97
- Davy, D. (2015). Understanding the Support Needs of Human-Trafficking Victims: A Review of Three Human-Trafficking Program Evaluations. Journal of Human Trafficking, 1(4), 318-337. DOI: 10.1080/23322705.2015.1090865
- EASO (2018). Briefing Paper Access to Information in the Context of Asylum: Exploring Existing Resources, Good Practices and Ways Forward. Available at: https://www.easo.europa.eu/sites/default/files/EASO-Briefing-Paper-Access-to Information--in-the-context-of-Asylum.pdf
- Eldridge, S., & Lancaster, G., & Campbell, M., & Thabane, L., & Hopewell, S., & Coleman, C., & Bond, C. (2016). Defining Feasibility and Pilot Studies in Preparation for Randomised Controlled Trials: Development of a Conceptual Framework. PLoS ONE, 11(3). 10.1371/journal.pone.0150205.
- Evans, W.N., & Perry, B., & Factor, R. (2019) The Impact of Broadband Access on Social and Economic Integration: Results from the RefugeeMobile Pilot. *Journal of Refugee Studies*. Published online: 26 August 2019. https://doi.org/10.1093/jrs/fez067
- EWSI (2020) European Web Site on Integration: Migrant integration information and good practices database. Available at: https://ec.europa.eu/migrant-integration/main-menu/eus-work/archive/actions
- Ezard, N., & Debakre, A., & Catillon, R. (2010). Screening and brief intervention for high-risk alcohol use in Mae La refugee camp, Thailand: A pilot project on the feasibility of training and implementation. *Intervention: International Journal of Mental Health, Psychosocial Work & Counselling in Areas of Armed Conflict*, 8(3), 223–232. https://doi.org/10.1097/WTF.0b013e3283413134

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 822688.

- Helbling, M., & Bjerre, L., & Römer, F., & Zobel, M. (2017). Measuring immigration policies: The IMPIC database. European *Political Science*, 16(1), 79-98.
- Huaibo, X., & Bailey, R., & Jiang, W., & Aronson, R., & Strack, R. (2011). A Pilot Intervention for Promoting Multiethnic Adult Refugee Groups' Mental Health: A Descriptive Article. *Journal of Immigrant and Refugee Studies*, 9 (3), 291-303.
- Jacobson, K., & Marshak, A., & Ofori-Adjei, A., & Kembabazi, J. (2006). IDP LIVELIHOODS: Using Microenterprise Interventions to Support the Livelihoods of Forcibly Displaced People: The Impact of a Microcredit Program in Idp Camps in Lira, Northern Uganda. *Refugee Survey Quarterly*, 25 (2), 23-29.
- Kaiser, T. (2004). Participation or Consultation? Reflections on a 'Beneficiary Based' Evaluation of UNHCR's Programme for Sierra Leonean and Liberian Refugees in Guinea, June–July 2000. *Journal of Refugee Studies*, 17(2), 185-204.
- Kivijärvi, A., & Aaltonen, S., & Välimäki, V. (2019). Feasibility of an online discussion group as a component of targeted youth work in Finland. Children and Youth Services Review. Published online 4 July 2019.
- Knowles, L.M. & Stelzer, E.-M. & Jovel, K.S. & O'Connor, M.-F. (2017) A pilot study of virtual support for grief: Feasibility, acceptability and preliminary outcomes. Computers in Human Behavior, 73, 650-658, [10.1016/j.chb.2017.04.005](10.1016/j.chb.2017.04.005)
- Kupzyk, Sara S., & Banks, Brea M., & Chadwell, Mindy R. (2016) Collaborating with Refugee Families to Increase Early Literacy Opportunities: a Pilot Investigation. Contemporary School Psychology, 20, 205–217.
- Lipscomb, M. (2008) Mixed method nursing studies: A critical realist critique. Nursing Philosophy, 9, 32-45.
- Owen, R., & Macnaghten, P., & Stilgoe, J. (2012). Responsible research and innovation – From science in society to science for society, with society. Science and Public Policy, 39(6), 751-760.
- Patton, M. Q. (2014). Qualitative inquiry in utilization-focused evaluation. In L. Goodyear, E. Barela, J. Jewiss & J. Usinger (eds.) Qualitative inquiry in evaluation: From theory to practice (pp. 25–54). San Francisco: Jossey-Bass.
- Pawson, R. (2013). *The Science of Evaluation. A Realist Manifesto*. London: Sage.
- Pawson, R. & Tilley (1997). *Realistic Evaluation*. London: Sage.
- Piwowarczyk, L., & Kelley, H.B., & Sondra, S., & Francine, C., & Mudymba, T., & Hashi, I.M., & Raj, A. (2013) Pilot Evaluation of a Health Promotion Program for African Immigrant and Refugee Women: The UJAMBO Program. *Journal of Immigrant and Minority Health*, 15 (1), 219–223.
- Potocky, M. (2010). Effectiveness of services for victims of international human trafficking: An exploratory evaluation. *Journal of Immigrant & Refugee Studies*, 8(4), 359–385. doi:10.1080/15562948.2010.522462
- Robson, Colin (2000). *Small-Scale Evaluation*. London: Sage.
- Robson, Colin & McCartan, Kieran (2016). *Real World Research. Fourth Edition*. London: Wiley.

- Sayedhabibollah A., & Izadikhah, Z., & Renzaho A., & O'Connor, P. (2019). Effectiveness of Psychological Interventions on Young Refugees' Social Adjustment: A Meta-analysis. *Journal of Refugee Studies*. Published Online: 6th May 2019, https://doi.org/10.1093/jrs/fez035
- Sayer, Andrew (2000). *Realism and Social Science*. London: Sage.
- Scriven, M. (1967). The Methodology of Evaluation. *AERA Monograph Series in Curriculum Evaluation*, 1, 39-83.
- Sen, A. (1992). *Inequality re-examined*. Oxford: Clarendon Press.
- Sen, A. (1993). Capability and Well-being. In: M. Nussbaum and A. Sen, eds., *The Quality of Life*. Oxford: Clarendon Press. pp. 30–53.
- Sekhon, M. & Cartwright, M. & Francis, J.J (2017) Acceptability of healthcare interventions: An overview of reviews and development of a theoretical framework. *BMC Health Services Research*, 17(88), 10.1186/s12913-017-2031-8
- Shadish, W. R., & Cook, T. D., & Campbell, D. T. (2002). Experimental and quasi-experimental designs for generalized causal inference. Houghton: Mifflin and Company.
- Stange, K. C., & Phillips, W. R. (2007). In This Issue: Real Change Is Real Hard in the Real World. *Annals of Family Medicine*, *5*(2), 98–100. doi:10.1370/afm.684
- Stringer, E. T (2013). *Action Research*. London: Sage.
- Tickle-Degnen, L. (2013). Nuts and Bolts of Conducting Feasibility Studies. *The American Journal of Occupational Therapy*, 67, 171-176. 10.5014/ajot.2013.006270.
- Turrini G. et al. (2019). Efficacy and acceptability of psychosocial interventions in asylum seekers and refugees: systematic review and meta-analysis. *Epidemiology and Psychiatric Sciences*, 28(4), 376–388.
- Vedung, E. (2010). Four waves of evaluation diffusion. *Evaluation*, (16), 263–277.
- Veenhoven, R. (2008). Sociological theories of subjective well-being. In M. Eid & R. Larsen (eds.) The Science of Subjective Well-being: A tribute to Ed Diener, pp. 44-61. New York: Guilford Publications.
- Webster, M. & Sell, J. (eds.) (2017). *Laboratory Experiments in the Social Sciences*. London: Elsevier.
- Wieland, M. L., & Njeru, J. W., & Hanza, M. M., & Boehm, D. H., & Singh, D., & Yawn, B. P., & Sia, I. G. (2017). Pilot Feasibility Study of a Digital Storytelling Intervention for Immigrant and Refugee Adults With Diabetes. *The Diabetes Educator*, *43*(4), 349–359. https://doi.org/10.1177/0145721717713317
- UNHCR (2020). *Research and Evaluation*. Available at: https://www.unhcr.org/evaluation-and-research.html
- Yankey, T., & Biswas, U. N. (2012). Life Skills Training as an Effective Intervention Strategy to Reduce Stress among Tibetan Refugee Adolescents. *Journal of Refugee Studies*, 25(4), 514–536. https://doi.org/10.1093/jrs/fer056

Coordinator contact: Dr. Rubén Fuentes-Fernández | Universidad Complutense de Madrid | Calle del Profesor José García Santesmases, 9. Ciudad Universitaria 28040 MADRID, Spain.
t: +34/ 91 394 7548 | e: rfuentes@ucm.es | w: www.ucm.es , grasia.fdi.ucm.es

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 822688.

Annex 1: A detailed list of actor-oriented criteria

PROCESS criterion	Frequency	Review	Reason if excluded	Evaluation category (if included)
Recognising/respecting FDP diversity	5	Included		Accessibility/acceptability
Respecting/listening FDPs	5	Included		Accessibility/empowerment
Recognising gender	4	Excluded	Included in criterion on diversity	Accessibility/acceptability
Respecting/listening the service providers	4	Included		Viability/empowerment
Individualised/group specific service provision	3	Excluded	Included in criterion on diversity	
Fostering cross-professionalism and cross-sectionalism	3	Included		Viability
Sensitivity to local context / operational malleability	3	Included		Viability
Recognising FDP skills	2	Excluded	Included in criterion on diversity	
Accessibility /wide reach	2	Included		Accessibility
Long enough processes	2	Excluded	Included in other criteria	
Anti-discrimination	2	Excluded	Included in criterion on diversity	
Efficiency in management	1	Excluded	Included in criterion on service providers	
Focusing on people	1	Excluded	All pilots should match before implementation	

Coordinator contact: Dr. Rubén Fuentes-Fernández | Universidad Complutense de Madrid | Calle del Profesor José García Santesmases, 9. Ciudad Universitaria 28040 MADRID, Spain.
t: +34/ 91 394 7548 | e: rfuentes@ucm.es | w: www.ucm.es , grasia.fdi.ucm.es

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 822688.

OUTCOME criterion	Frequency	Review	Reason if excluded	
Increase autonomy / knowledge / participation	9	Included		Wellbeing/inclusion
Improving living conditions / material wellbeing	5	Included		Wellbeing
Access to services/education	5	Included		Inclusion
Promoting cohesion over group boundaries	5	Included		Inclusion
Foster professional skills/awareness	5	Included		Capacity/inclusion
Increase awareness on labour market / employment rate	4	Excluded	Included in criterion on professional skills etc.	
Improve language skills	3	Included		Capacity
Foster empowerment	3	Excluded	Included in criterion on autonomy etc.	
Enhancing emotional/psychological support	3	Excluded	Included in criterion on access to services etc.	
Increase knowledge on rights / available services	2	Included		Capacity
Increase cross-sectoral / professional co-operation	1	Included		Competence
Reducing hate-speech	1	Excluded	Not targeted & difficult to assess	
Changing/affecting legislation	1	Excluded	Out of scope of project time limits	
Facilitate FDPs adjustment in integration policies	1	Excluded	Not targeted	
Foster engagement in one's own community	1	Excluded	Not targeted	
Increase the number / accessibility of voluntary workers	1	Excluded	Not targeted	
Promoting the registration of FDPs	1	Excluded	Not targeted	
Preventing substance abuse	1	Excluded	Not targeted	

Coordinator contact: Dr. Rubén Fuentes-Fernández | Universidad Complutense de Madrid | Calle del Profesor José García Santesmases, 9. Ciudad Universitaria 28040 MADRID, Spain.
t: +34/ 91 394 7548 | e: rfuentes@ucm.es | w: www.ucm.es , grasia.fdi.ucm.es

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 822688.