

## Accepted Manuscript

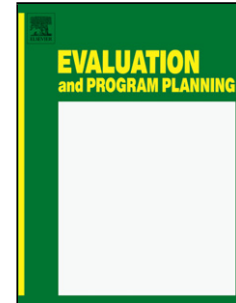
Title: Evaluability assessment of a community development leisure program in Spain

Authors: María Hortigüela Arroyo, Silvia Ubillos Landa

PII: S0149-7189(17)30042-3

DOI: <https://doi.org/10.1016/j.evalprogplan.2018.10.014>

Reference: EPP 1614



To appear in:

Received date: 10-2-2017

Revised date: 7-9-2018

Accepted date: 9-10-2018

Please cite this article as: Arroyo MH, Landa SU, Evaluability assessment of a community development leisure program in Spain, *Evaluation and Program Planning* (2018), <https://doi.org/10.1016/j.evalprogplan.2018.10.014>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

**Running Header: Evaluability assessment of a community development leisure program in Spain****Information about the authors:**

María Hortigüela Arroyo earned her PhD in social pedagogy at the University of Burgos (Spain). Her research activities currently focus on the assessment of social programs, subjective quality of life and community development programs.

E-mail: marhorti@hotmail.com

Silvia Ubillos Landa earned her PhD in psychology from the University of the Basque Country (Spain) and currently works as a professor at the University of Burgos Department of Health Sciences. Her research activities focus mainly on the assessment of community, educational, social and health programs and the development of evaluation instruments. She is the author of numerous publications in this particular field of research, and has also given many lectures, presentations and courses on program assessment.

E-mail: subillos@ubu.es

Corresponding Author:

Silvia Ubillos Landa

Facultad de Humanidades y Educación, C/ Villadiego s/n, 09001 Burgos. E-mail:

[subillos@ubu.es](mailto:subillos@ubu.es) Tel.: 947 25 87 23

**Highlights**

- A qualitative validation of a theoretical evaluability assessment model is proposed.
- Practical tools for reliably conducting an evaluability assessment are provided.
- Evaluability assessment ensures that future evaluation undertaken will be effective.

**1. Introduction**

Although evaluability assessments (EAs) were originally considered a tool or activity used prior to an evaluation (Thurston & Potvin, 2003), in current practice they have evolved into an

evaluation approach that can be used at any point of a program's lifecycle and can serve formative, summative and developmental evaluation purposes (Trevisan & Walser, 2015).

Since the concept was first developed by Wholey during the nineteen-seventies (1979), EA has evolved substantially, passing through a number of different phases, each characterized by a specific vision and orientation. Thus, after the nineteen-seventies, in which the concept focused on analyzing the structure of a program in order to determine its evaluability, there came a phase characterized by the clarification and methodological systematization of EA, which reached its peak during the nineteen-eighties with the emergence of multiple models which expanded on the original approaches outlined by Wholey. Two of the most commonly cited of these models are the ones by Rutman (1980) and Smith (1989). Parallel to this methodological enrichment, the approach itself also underwent substantial change, with the focus shifting towards defining the detailed theory underpinning a program (Hurtworth, 2008) and determining the feasibility of implementing the necessary evaluation (Trevisan & Walser, 2015). Following this period, and up until 2007, there was a decrease in the use of EAs, with some of the reasons offered being: (a) EAs often went unreported (Smith, 2005) and reviews of their use were scarce (Davies, 2013); (b) the development of program theory and the sophistication of program logic models as evaluation strategies in their own right (Hurtworth, 2008); (c) the realization that program development, implementation, and evaluation do not necessarily flow in an orderly fashion (Trevisan 2007); (d) lack of accessible guidance materials and a clearly defined methodology (Davies & Payne, 2015) and (e) ambiguous articulation of EA as a concept.

It was in this context that our EA model arose as a means of offering a solution to the shortcomings outlined in the paragraph above. The research carried out was not based on a negation of the validity of classic EA models, but rather on the realization that there are simply too many of them, along with a lack of clarity regarding their application. We therefore proposed a model which combined the strengths of the 3 principal models identified in our review of the published literature, thereby compensating for the limitations of each. We labeled the models identified the *European conceptual model*, the *Spanish check-list* and the *American pragmatic model*, a categorization which is consistent with the subsequent review carried out by Davies (2013), who also identified 3 main EA models.

The results obtained are timely, as we are witnessing a resurgence of EAs (Trevisan & Walser, 2015), and their use among agencies and organizations (Craig & Campbell, 2015; D'Ostie-

Racines, Dagenais & Ridde, 2013). This revival is partly due to demands from governments sponsoring programs, who have found EAs useful for ensuring good governance, transparency and accountability for public actions (Davies & Payne, 2015), increasing the effectiveness of the earmarked budgets, generating spaces of trust and helping strengthen the culture of evaluation (Craig & Campbell, 2015). Within this trend, results have recently been published regarding the application of EAs to national social-healthcare (Aquilino, Arias, Estévez & Suaya, 2015) and healthcare programs (Belford, Robertson & Jepson, 2017). However, according to our review of the literature and knowledge to date, this is the first EA carried out of a community development leisure program in Spain.

The present study therefore helps further existing literature on community development EAs and constitutes a novel contribution in the Spanish context, in which their use, as March, Orte and Ballester (2016) state, is not generalized. Moreover, our aim was to ensure that, in addition to being of use for the assessed program itself and the stakeholders involved in it, the dissemination of the positive results obtained in this study would generate pedagogical value at a scientific level, thus increasing both the use and credibility of EAs in general.

In specific terms, our study aims to contribute to the literature on EA, considering both scholars' and practitioners' needs and preferences, as it encompasses some of the characteristics proposed by Trevisan (2007) for fostering the use of EAs:

(a) *It is based on the need to organize and present EA in a more simplified manner than in classic EA models.* Since, in practice, EA is not a linear sequence of discrete steps (Trevisan & Walser, 2015), instead of steps, our model is comprised of three main components that are common to all EAs, thereby enabling its application in different contexts. Moreover, it is presented graphically, with the aim of rendering it more user-friendly while still maintaining the essential elements of EA.

(b) *The original nature of the context in which the EA was carried out.* Our goal was to create a robust and innovative EA model for use in the fields of public administration and community development. To this end, in addition to including the more widely-cited models in existence, we also incorporated less well-known authors whom we nevertheless consider to be relevant in the field of community development and the administration of public policies, such as those included in the European conceptual model and the Spanish check-list.

(c) *It compares the usefulness of various EA models by creating a holistic model.* Our proposal compensates for the shortcomings of historical models by attaching greater importance to the construction of the theoretical model and adding specific methodologies which, in our case, can be replicated in any context, but are specifically oriented towards the public administration.

(d) *It integrates EA into a broader evaluation plan, i.e. an evaluation of results which was carried out as part of another study.* In our EA model, we applied qualitative methodology with a generative aim, which is linked to the development of theories, models, strategies and actions (Ritchie, 2003), and in this case is associated with a model applied within the context of the public administration.

(e) *It includes standards of quality evaluation and outlines how they can be applied to the performance of an EA.* Qualitative research takes into account the fact that viewpoints and practices in the field are different due to the different subjective perspectives and social backgrounds related to them (Flick, 2014). In order to render this type of qualitative methodology both rigorous and valid, we based our study on the model of trustworthiness of qualitative research proposed by Lincoln and Guba.

In specific terms, the research aims were as follows: firstly, to design, apply and qualitatively validate a comprehensive EA model; secondly, to clarify the theoretical model of the program, thirdly, to analyze the implementation of the program and determine whether or not the quality of its planning and execution enabled an in-depth evaluation; and finally, to select the evaluation design most suitable for the program, according to the needs, barriers and circumstances detected in the EA.

## **2. Background**

The EA was conducted in Ibeas de Juarros, a small rural town with 1345 inhabitants, located in Spain. Over recent decades, the town has undergone an unexpected population increase, due to the arrival of young families attracted by lower real estate prices. This, in turn, has disrupted the social cohesion which existed among the original inhabitants, since newcomers do not have the same ties to the local region and do not tend to participate actively in the community. In light of this situation, the Local Council decided to take action to encourage a greater degree of

integration among younger members of the population who did not play an active role in the community. The approach adopted was to make the town's civic center more dynamic, with the aim of improving the quality of life of all inhabitants and achieving a greater degree of cohesion among the community. To this end, the Council initiated a Community Development Program (CDP) aimed at all age groups and segments of the population. The CDP is directed and funded by the Ibeas de Juarros Local Council, through its Department of Culture, although other municipal areas are also involved in its implementation (Sports, Social Services and Rural Development Departments). The CDP is carried out in the Ibeas de Juarros civic center, which is a municipal outreach building that provides educational, training, cultural, social, citizen support and sports services with a fair degree of coordination and fosters both closer relations among residents and citizen engagement. The CDP comprises four main programs:

*Social promotion program for adults:* Its aim is to improve well-being and personal development through the acquisition of personal skills, cultural knowledge and the promotion of physical activity.

*Program for families and children:* It focuses on children and aims to help parents achieve a better work-life balance through non-formal education and leisure services.

*Stimulation and engagement program:* It fosters the development of social networks through support for associations, the free provision of municipal areas and premises, awareness-raising and the organization of solidarity events.

*Rural development and tourism program:* It includes cultural activities which aim to raise awareness of the area's inherent tourism, gastronomic and agricultural resources.

In addition to the four programs outlined above, the CDP also provides three free services which serve to improve users' biopsychosocial well-being and enhance social cohesion and integration among citizens. These services are: *Municipal library*, *Information and citizen support service* and *Telecenter service*.

Nine years after the creation of the CDP, a decision was made to evaluate the results achieved so far. Given that it had never previously been assessed, a prior EA was recommended.

### **3. Methods**

#### **3.1. Development of the EA model**

The research study started with an exhaustive review of the literature on EA published to date (2012), from classic models (e.g.: Alvira, 1991; Rossi et al., 2004; Rutman, 1980; Smith, 1989; Thurston & Potvin, 2003; Vedung, 1997; Wholey, 1979) to more recent practical applications (D'Ostie-Racines et al., 2013; Reiman, 2012). The review led to the conclusion that, as Davies and Payne (2015) state, at that moment in time there was a lack of clearly defined and easily accessible methods. Moreover, in the field of public program evaluation, which is the field in which our study is located, although some theoretical models of EA were found, they lacked either validation (e.g., Alvira, 1991; Merino, 2007) or methodological development (e.g., Vedung, 1997). The review also found some isolated cases of methodologies with no allusion to any theoretical model (e.g., Dunn, 2008; Fernández-Ballesteros; 1995; Kaufman-Levy & Poulin, 2003; Reiman, 2012). The decision was therefore made to create our own model which would encompass those proposals most applicable to the public evaluation of programs and which, in the interests of conceptual clarity, we have summarized and divided into categories under the following headings (presented here in chronological order):

*Model 1. European conceptual model:* Represented by two pioneers in the adaptation of the generalist American model to the European context of public programs: Alvira (1991) and Vedung (1997). This model is characterized by its focus on identifying the program's theoretical model, conducting a future evaluation feasibility analysis and determining the improvements required to enable the program to be evaluated effectively.

*Model 2. Spanish check-list:* As check-lists are one of the most recommended and widely-used tools in EAs (Davies, 2013), we decided to include under this heading the Spanish approach based on the work of Fernández-Ballesteros (1995), who developed a methodological tool in the form of a list of relevant questions. Of special interest to our research is the innovative adaptation of the tool to public programs and policies, carried out by the Navarre Institute of Public Administration (NIPA, 2008). However, this second model contains no sequenced global model like the American model, and nor does it include an explanatory model like its European counterpart.

*Model 3. American pragmatic model:* This model encompasses proposals focused on concisely-described phases which nevertheless make no explicit reference to specific theoretical models (unlike in the first model). The proposals include lists of key questions for the evaluator, to help him or her conduct semi-structured interviews with stakeholders. They also include check-list templates and the most suitable design for future evaluation. Some of the most cited authors in

this model are, for example, Dunn (2008), Kaufman-Levy and Poulin (2003) and Thurston and Potvin (2003).

After analyzing all three models and identifying their strong points and shortcomings, a new, comprehensive, evaluability design was developed (figure 1):

### 3.2. Phases of the EA model

Our model comprises three consecutive phases focusing on three issues which, according to the review by Davies (2013), most authors agree an EA should consider: Program Design, Availability of Information and Institutional Context. Moreover, the actions to be taken in each phase have been identified and the most appropriate methodological tools for each have been selected. The execution of each phase results in an output that is sequenced in relation to the previous one.

*Phase 1. Theoretical Model.* This phase examines the adequacy of the program design. The actions that must be undertaken are: 1.1. Theoretical Description: this includes preparing an overview of the program based on official material, administrative material and normative instructions (among others) in order to outline its components, goals or expected effects. The methodological tool we propose is the analysis of formal documents. 1.2. Extraction of the Theory: this refers to determining stakeholder perceptions of the program, using semi-structured interview guidelines. 1.3. Evaluation of the Theory: the main purpose here is to chart the flow that the program should follow to produce its effects, in accordance with the document analysis. In situ observation of the programming process is suggested to achieve this goal. Once this phase has been completed, the Program Theory will be clear, and phase 2 will be carried out on the basis of this output.

*Phase 2. Evaluable Model.* This phase focuses on the evaluability of the program in practice, given the availability of data required to carry out an evaluation and the systems able to provide it. The actions to achieve in this phase are: 2.1. Evaluation of the program: this involves gathering more information in the field in order to gain a broad vision of how the program actually works. Direct, non-participatory observation of the program is suggested at this point. 2.2. Development of the evaluable model: this stage determines which components and goals/effects of the program meet the preconditions for being evaluated. The process also enables you to verify whether or not the information obtained is suited to a standardization of



procedures that allows a mature evaluation of the program to be carried out. Useful methodological tools that can be applied in this action are check-lists, semi-structured interviews, focus groups, surveys and the Delphi technique, depending on the characteristics of the program. The output of this phase will be the Program Logic Model, a document that will be generated by gathering all the information, studying it and comparing the material with all stakeholders.

Phase 3. *Model to be Evaluated*. This phase focuses on the conduciveness of the context to evaluation. The main action is the drafting of the Evaluability Report, a document that summarizes, collates and analyzes all the data of the phases within the EA, justifies whether or not it is worthwhile carrying out a future evaluation and lays out possible alternatives and recommendations regarding the evaluability of the program. The third phase culminates in the generation of the Evaluation Plan, in which the evaluation users are identified and the possibilities of developing a full-scale evaluation based on the data obtained with the EA are discussed with them. Semi-Structured interviews are the most useful tools at this stage.

### 3.3. Characteristics of the model

The main characteristics of the model created are:

*Integrating*: It integrates all three EA approaches described above, compensating the lack of a theoretical basis found in both models 2 and 3 with the extensive theoretical development of model 1. In order to integrate all three proposals into a single model, from the European approach we used Vedung's proposal (1997) as the general guideline for the model, carrying out an analysis of the program in order to construct first the Theoretical Model, then the Evaluable Model and finally the Model to be Evaluated. From the Spanish check-list model we adapted the tool developed by the NIPA (2008) to create a simplified version that was easier to understand and apply in the context in which we were going to be conducting the EA. Finally, from the American pragmatic model, we took the most significant characteristics, namely clarity, concise description of the different phases proposed by each author for carrying out an EA and the useful interview guidelines for the different stakeholders.

*Holistic*: In our proposal, the global model which came out of the sum of the different approaches gave rise to a more comprehensive and global EA than would have resulted from an isolated analysis of the program's theory, the application of sequenced phases with no underpinning conceptual model, or the mere use of one or more check-lists as a stand-alone tool

with no general explanatory model. Furthermore, we also believe that it resolves one of the main criticisms leveled at this type of assessment, namely the disjointed use of actions peripheral to the evaluation, since it offers a holistic EA model based on a uniform rationale which underpins and justifies the successive evaluation actions carried out.

*Flexible:* The methodological design is established as the EA progresses, remaining open and flexible to any changes or redefinitions that may be required either to the assessment itself or to the evaluation strategies used to conduct it. It is therefore able to respond and adapt to different situations. This characteristic does not imply that any one of the proposed steps can be omitted. Rather, it means they can be reordered as required, mainly in accordance with the suitability of application times. The model's flexibility also alludes to the methodological tools proposed, which must be adapted to the characteristics of the specific program being assessed. Thus, circumstances may require, for example, the use of structured interviews (if the size of the program renders this advisable), focus groups (in the case of larger programs), the Delphi technique (in the case of very large programs) or even a survey of the target population (in the case of programs implemented in one or more countries at the same time), without any of these detracting in any way from the validity of the general proposal. In all cases, however, the vital instruments required to conduct an EA are: a document analysis and stakeholder interviews (to construct the Theoretical Model) and the use of check-lists and direct observation (to develop the Evaluable Model). In short, our proposal integrates the main approaches in the field of EA in order to offer a holistic, pedagogic instrument which enables the use of a flexible methodology adapted to different contexts, thus fostering transferability.

*Rigorous:* The proposed model is based on the premise of scientific rigor and is therefore subject to quality controls in order to guarantee its validity. Given that the characteristics of the object of the evaluability study require the use of an eminently qualitative methodology, the quality control applied to the model is necessarily suited to that methodology. Specifically, Lincoln and Guba's excellence criteria (1985) were used. These guidelines guarantee the trustworthiness of the results obtained using the qualitative methodology in accordance with the following criteria: transferability instead of external validity (the focus is more on the application of the results); credibility instead of internal validity (more focus on the true value of the research); dependability instead of reliability (more focus on the consistency of the results) and confirmability rather than objectivity (in reference to the problem of neutrality). The techniques used to ensure the trustworthiness of the model in each of the criteria are explained in section 5.4.

#### 4. Procedures

To verify the practicality of the EA model and its feasibility, a qualitative study was conducted on the Ibeas de Juarros CDP over the course of 9 months. The process commenced with the construction of the Theoretical Model (phase 1) and the Theoretical description of the program (1.1) through the document analysis of internal CDP texts (activity reports, population data, etc.) and current legislation. The ultimate aim was to contextualize and understand: the background, the problem to be solved, the institutional context, funding sources, implementation levels, coverage and specific actions taken. Having analyzed the information, contact was made and semi-structured interviews were conducted with the CDP stakeholders, in order to achieve the Extraction of the Theory (1.2). Seven one-hour semi-structured interviews were conducted and recorded, using scripts designed in accordance with the specific stakeholders in question, adapted from the third model.

This information was later complemented by direct, non-participatory observation of a general programming meeting. Throughout the Evaluation of the Theory (1.3), all the most relevant information was analyzed, interpreted and integrated. Different drafts of the Program Theory were proposed and discussed by the stakeholders directly responsible for the CDP, until a definitive agreement was reached regarding the theoretical basis upon which the program rested, and a general goal or objective was drawn up.

In order to develop the Evaluable Model (phase 2) and the Model to be Evaluated (phase 3), an in-depth analysis was conducted of the principal dimensions affecting the evaluability of the CDP. First, the stage named *Evaluation of the program* (2.1) was analyzed through the information gathered about the way the program operates from the 8 in situ observation sessions, which covered the entire process, from the prior coordination right up to the dissemination, registration, human, material and economic resource management and the running of the activities themselves. Next, the *Development of the Evaluable Model* (2.2) stage was implemented, asking technical experts and politicians to complete a check-list. The methodological tool applied was our own adaptation of the protocol developed by the NIPA (2008). The final instrument comprised 49 items, grouped into: Planning quality (18 items), Quality of the information system (12 items), Clarity of the evaluation aims (7 items), Resources earmarked for evaluation (6 items) and Barriers associated with the evaluation (6 items). The response format was dichotomous (yes/no). Finally, this information was further complemented by an analysis of the perceptions of both CDP beneficiaries and non-beneficiaries, gathered

through semi-structured interviews with the pertinent stakeholders, following the guidelines adapted from the pragmatic American model. The outputs of the second (Logic Model) and third (Evaluation Plan) phases are discussed in the results section.

## 5. Results

The results obtained were reflected in three documents that the CDP lacked prior to the EA, and which mainly serve to aid the program's future planning, development and evaluation. These documents are:

### 5.1. Program Theory

The EA process identified a clear lack of a formal theoretical model for the CDP, understood as a formally approved programmatic written document. This circumstance is not unusual for CDPs implemented in rural contexts, where work is mainly carried out reactively, providing dynamic responses to needs as and when they arise. Stakeholders involved in the programming phase are asked to come up with proposals and activities in very short periods of time, in order to provide evidence that the initiative is making headway. This working method helps justify the use of the earmarked budget but overlooks the importance of reasoned planning. Furthermore, although theoretically speaking, correct program planning includes the operationalization of the aims, thus enabling a final assessment of the degree of efficacy and efficiency reached in their achievement, in practice, in social or leisure community development programs it is more usual, as March et al. (2016) claim, to find designs that are far from rigorous as regards the definition and precise identification of goals, as well as a complete absence of evaluative processes in the general planning of the program or service. In this study, the main outcome of the EA was the extraction of the Program Theory, based on a set of diffuse intentions and disjointed actions. In specific terms, the study resulted in a 10-page programmatic document which was drafted in a participatory manner and was then ratified by the Local Council as the basis for the general planning of the CDP. The document includes: legal and theoretical justification of the problem and the needs to which the CDP aims to respond, and a consensus-based definition of the general aim of the program, to serve as the basis for future activities.

### 5.2. Logic Model

The second document generated is linked to the Evaluable Model, with the main by-product being a proposal for a Logic Model in the form of a flow chart to describe the CDP (figure 2).

Nevertheless, it is important to point out that the use of logic models is limited unless all the stakeholders in the program participate in the development process. Consequently, a complete logic model was not developed during the course of the EA, since it was not one of our research aims to intervene in the establishment of specific aims, indicators or achievement measures for each CDP service or activity. We believe this is a technical and political task that requires a long period of time and which exceeds the evaluative competences of an EA. Thus, our contribution to the logic model consisted of the triangulation and synthesis of the information gathered through the administration of the check-list to technical experts and politicians, direct non-participatory observation of program implementation and the semi-structured interviews with beneficiaries and non-beneficiaries of the CDP, reflected in the flow chart. This process serves to explain, in a reasoned manner, the structure of the CDP programs and services, and helps ensure compliance with its general aim. This proposal of a logic model may offer an excellent starting point for the team of technical experts and politicians, who may further enrich it by including the specific aims they wish to achieve through each program and service, the indicators they plan to use to measure the degree to which targets are met, the instruments or techniques required to carry out this measurement and the technical, economic and human resources they plan to earmark for program evaluation.

### *5.3. Evaluation Plan*

The third document which resulted from the EA was an Evaluation Plan based on the evaluability report, which concluded that, at present, the CDP is not ready to admit an internal results evaluation. Rather, it is still at a prior stage, namely the planning stage. As indicated above, the lack of systematic planning was the main obstacle detected for conducting an evaluation based on target compliance, which is usually considered a summative assessment of the program, i.e. an ex-post evaluation, carried out internally and focusing on the results of the intervention, measured through the program's outcome products. Among evaluators who use EAs there are a number of different stances regarding the activist role such evaluations can play in programs (Vedung, 1997). Some authors use EAs to restructure programs in order to render them evaluable (non-naturalist evaluation strategy), while others use them with no ambition to

adapt the target programs to the needs of a hypothetical future evaluation (naturalist evaluation strategy). An intermediate approach is that postulated by Rutman (1980), who believes that the EA process may trigger changes that enhance the program's evaluability, and that said changes can be considered by-products of the EA, rather than its main purpose. This intermediate approach was the one used in this study. A naturalist strategy was adopted as the main aim of the EA, in order to determine whether or not the CDP could be evaluated. The conclusion was that, at this moment in time, it was not feasible (either technically or economically) to carry out an internal program evaluation. Nevertheless, an evaluative design adapted to the current situation of the CDP was proposed. The recommendation was an external ex-post evaluation of the results, using a quantitative method and an experimental pre-test/post-test repeated measurement design with an equivalent control group, in order to measure the impact of the CDP on participants' subjective quality of life and social cohesion in comparison with that of non-participants. Moreover, a series of changes were suggested in the evaluability report which, if introduced into the CDP planning process, would guarantee its future evaluability. These changes referred mainly to: the Quality of planning, the Quality of the information system, the Clarity of the evaluation aims, the Specification of the resources earmarked for evaluation and the Overcoming of barriers linked to evaluation.

#### 5.4. *Qualitative validity of the study*

While quantitative researchers consider *reliability*, *objectivity* and *validity* to ensure the trustworthiness of the findings, Lincoln and Guba's (1985) trustworthiness criteria are the most commonly used (Anney, 2014; Ruiz- Olabuénaga, 2012) to ensure the rigor of qualitative findings. The following techniques were used to assess each of the four trustworthiness criteria of the EA model:

*Transferability*: refers to the degree to which the results can be transferred to other contexts. The main strategies used are thick description and purposive sampling, as when the researcher provides a detailed description of the enquiry and explains why participants were purposively selected, this facilitates the transferability of the inquiry research (Anney, 2014). In our study, this criterion was particularly important during the construction of the Program Theory and was guaranteed through the use of the thick description method, providing detailed, dense descriptions of all the steps carried out, thus establishing a reasoned basis for likeness judgments. However, above all, this criterion strongly influenced the selection of the EA

participants, a process during which the purposive sampling technique based on intentional criteria (Ruiz-Olabuénaga, 2012) was used, since transferability is achieved not in accordance with the number of subjects analyzed, but rather in accordance with their particular knowledge of the issues under investigation (Anney, 2014). Specifically, we applied strategic criteria by selecting representative informants from all stakeholder groups identified by Greene (2005):

- People with the authority or capacity to make decisions about the program, including politicians, sponsors and consultants, etc. The political representative responsible for the program was interviewed in May 2013.
- Those directly responsible for the implementation of the program, such as technical experts, managers, administrators, executives and agents, etc. Two technical experts were interviewed in March 2013.
- Potential beneficiaries of the program, their family members or the communities of which they form part. Two participants in the CDP were interviewed between May and June 2013, and one member of an association collaborating in the CDP activities was interviewed in August 2013.
- Those who may potentially be harmed or inconvenienced by the program, or who have lost the opportunity to benefit from it. An interview was conducted with a local businessman in July 2013.

*Credibility*: establishes the confidence that research findings represent plausible information drawn from the participants' original data and is a correct interpretation of the participants' original views (Lincoln & Guba, 1985). It is a criterion which comes into play most strongly during the development of the Evaluable Model (although like the other criteria also, it was taken into consideration throughout the whole process, not just one single phase). Two specific techniques were used to guarantee it. Firstly, persistent observation which focuses intensely on those points of the program that are most characteristic. Secondly, triangulation of data collection techniques, comparing the information gained from the qualitative data and interpretations gathered from the members of the different stakeholder groups (Houghton, Casey, Shaw, & Murphy, 2013).

*Dependability*: deals with the stability of findings over time (Bitsch, 2005). In our EA, two strategies were used. The first was member control, involving stakeholders in the evaluation of findings and the interpretation of the recommendations of the study, in order to ensure that they are all supported by the data received from them (Anney, 2014). The second strategy was peer examination, in which the researchers discussed both the research process and the findings with

neutral colleagues. This strategy allows researchers to be honest (Bitsch, 2005) and contributes to deeper reflexive analysis (Anney, 2014) and was used throughout the EA.

*Confirmability*: refers to the degree to which the results can be corroborated by other researchers, to ensure that, as Tobin and Begley (2004) state: “data and interpretations of the findings are not figments of the inquirer’s imagination, but are clearly derived from the data” (p. 392). In our EA, the study was submitted to an audit trail analyzed by an external researcher, as a control process carried out to determine the consistency of the data and the relationship between the deductions made (Houghton et al., 2013). The researcher determined whether or not the research processes followed complied with accepted professional practice.

## **6. Lessons Learned**

The EA conducted of the CDP resulted in a series of recommendations regarding the model designed and the validation process itself, which we present here as a series of strong and weak points. These should be of use to evaluators initiating a similar process or seeking to replicate the one described here.

One of the model’s strong points is its pedagogic clarity and accessibility, which can be summed up in an easy-to-understand flow chart. This characteristic enables program stakeholders to quickly and intuitively understand and evaluate both the EA process itself and their participation in it.

Secondly, the model impacts the early stages of planning, the specification of the problem and the analysis of requirements, which are, in practice, seldom taken into consideration as key elements of the process. Its usefulness in the planning processes was also highlighted, since it serves as a guideline for establishing the minimum elements to bear in mind from the initial analysis of the situation onwards. The questions posed to determine and establish the three phases help identify the key aspects which characterize evaluable initiatives.

Thirdly, it is an easy-to-use model, since it does not require additional economic resources and is limited in time (approximately 6 months). These attributes enable evaluators to improve the quality of the target interventions in a short space of time, helping to systematically assess the extent to which the program is evaluable. In this particular study, the EA took longer than the established 6 months, since during the process a decision was made to actively contribute to the



consensus-based drafting and subsequent approval of the Program Theory, thus fulfilling the first two of what, according to Davies (2013), are the aims of any EA, namely “to improve the project design prior to approval; or to inform the design of a monitoring and evaluation framework in the inception period; or to decide if an evaluation should take place later on; or to inform the specific design of an evaluation that has now been planned for.”

Finally, one important result of the study was the success of the model for fostering personal reflection, group debate and self-assessment among reticent stakeholders. Its use is, therefore, recommended for all professionals and technical experts responsible for planning, implementation and evaluation processes who are interested in improving their practice and performance. Although during the study the model was tested by the authors, it was found that with some slight modifications it can also be used internally as a self-assessment tool by the staff running the program. Both the structure of the model itself and the instruments proposed foster a critical analysis of the actions and methodology employed by the teams responsible for a program. From this perspective, the model may serve as a catalyst for self-assessment and debates that help enhance and consolidate capacities. To achieve this, we suggest that the check-list be completed by at least two people related to the design, implementation and evaluation of the initiative, and that the conclusions drawn from each be compared and contrasted.

Our model also has a number of constraints and limitations. The following is an outline of the principal shortfalls identified and the various possibilities for improvement. Firstly, it is likely that the model proved easy to administer due to the research team's prior knowledge of EA and training in the field of evaluation. Nevertheless, it was observed that the complexity of the technical language used in the phases (e.g.: evaluable model, logic model, etc.) may limit the use of the instrument by some stakeholders who have no prior training in program evaluation (we refer mainly to politicians and staff). The improvement opportunities arising from this constraint focus on minimizing the prior knowledge of evaluation techniques required for using the instrument, an improvement which would also serve to enhance transferability. To this end, a brief user manual would need to be developed, outlining the profile of potential users, along with an instructions booklet designed to facilitate use by those with no prior training in evaluation or evaluability. As a result of its practical application, the model has gradually been improved and enriched and, at the same time, the check-list first developed by the NIPA (2008) has been adapted. While the basic structure of the check-list has been maintained, areas not previously covered, such as barriers to evaluation, have been included and the quantification

according to response percentages has been modified, since from an empirical perspective, it was difficult to quantify all the items in percentage terms. The response system was therefore replaced by a dichotomous response format. Another research limitation is that only one program was assessed, so we recommend future research to empirically validate the EA model in other contexts.

## 7. Discussion

The application of the EA had a strong positive impact on both the CDP itself and the stakeholders. Consequently, one of the principal conclusions that can be drawn in relation to the established aims is that the study validates the EA model proposed as a useful instrument for conducting an EA. Its usefulness is evident in its efficacy for both clarifying the Theoretical Model and analyzing the implementation of the CDP and how it fits with the Program Theory. In this sense, from a practical perspective, the most important result obtained in relation to the Program Theory was its identification on the basis of a set of indeterminate intentions and disjointed actions. In specific terms, working together with experts and politicians, the research team managed to achieve a greater degree of clarity in the program, providing legal and theoretical justification of the problem and needs to be addressed, and defining the CDP's general objective, which in turn may serve as a guideline for future programming and evaluation actions.

As regards the analysis of the program implementation and its fit with the Theoretical Model (the study's second aim), an Evaluable Model was developed which included a proposal for a logic operation model presented in the form of a flow chart of the four framework programs and three services that together make up the CDP. Therefore, the research project enabled the establishment of both the CDP Theory and its logic model, two key documents for conducting an evaluation, which were lacking in the program prior to the study. The conclusion regarding the third aim, i.e. to determine whether or not the quality of the planning and implementation process was sufficient to enable an evaluation, is that at the time the EA was conducted, neither the program itself nor its team (technical/political) was prepared to undertake a summative evaluation, since there was no systematic planning of sufficient quality on which to base this procedure. Nevertheless, the EA process did prove useful to the technical and political team, enabling them to:

- Obtain useful information for decision-making, thus helping them strengthen and perfect general programming and intervention processes (from needs analysis to evaluation).
- Gain a greater awareness of their current situation. In addition to the semi-structured interviews, the use of our check-list exemplified, in a very simple way, the protocol that should be followed when planning an evaluation that is cost-effective for the institution.
- Become aware of the need for and benefits of target-based action planning, rather than a reactive methodology that simply responds to problems and needs as and when they arise.
- Internalize the benefits of an evaluation, rendering the difference between measuring and evaluating clearer and more visible. The fact that the CDP's EA was conducted externally in a rigorous manner enabled the technical team to experience the evaluation procedure in an eminently practical way, thus helping to reinforce their performance.
- Improve the working climate among all those involved in the CDP, by raising awareness of the achievements made to date and recognizing the difficulties inherent in the coordination, management and planning of a program of this nature.

A non-interventionist stance was adopted throughout the entire EA process. Consequently, no changes were made to the program to improve the feasibility of a future evaluation.

Nevertheless, certain by-products were generated, namely a set of recommendations which, if adhered to in the future general planning process, would increase the program's evaluability. These recommendations refer to planning-related quality dimensions (mainly the definition of specific aims and indicators), the quality of the information system (basically the creation of such a system, since nothing was in place which would enable a systematic evaluation of the program), clarity of the evaluation proposal (stakeholders were encouraged to clarify the ultimate aim of the evaluation and the way the results would be used in the future), estimate of resources earmarked for evaluation (both human and economic, with the establishment of a new budget item for this) and the barriers associated with evaluation (since no barriers were detected, the recommendation was to strive to preserve this positive predisposition by creating a follow-up procedure of the evaluation which would enable all stakeholders to understand it and incorporate any suggestions made). Moreover, a series of strengths were detected for the CDP, which will enable stakeholders to continue to move along the path to effective evaluation. These strong points are:

- The program has scheduled a new general planning phase which will take into account the knowledge generated thanks to the EA. This is clearly beneficial for the technical and political

team, particularly if one accepts that the program's achievements and evaluability directly depend on the effort made during the design and implementation phases. In other words, if subsequent evaluation is taken into consideration right from the initial planning phase, it is more likely that said evaluation will be conducted in a much more rigorous, efficient and effective manner.

- No discrepancies were detected between the stakeholders as regards either the goals pursued by the program or the logical consistency of the program components (logic model). This indicates that the planning phase should be smooth and trouble-free, since it is merely a case of expressing in writing that which, to a certain extent at least, they are already doing, with just a few slight adjustments and improvements.
- There is a positive attitude within the organization towards evaluation, which was further enhanced by the EA, since stakeholders perceived it as enabling a higher level of rationality in the program management. The EA also helped by highlighting the need for evaluation and the benefits it would bring to the organization, technical experts and, above all, beneficiaries.

The final aim of the study was to select the evaluation design most suitable for the program, according to the needs, barriers and circumstances detected in the EA. After identifying the program's strengths and weaknesses as regards evaluation and the current situation of its planning, and taking into account the general aim agreed upon by consensus following the EA, the following feasible evaluation design was proposed: an external ex-post evaluation of the results, using a quantitative method and an experimental pre-test/post-test repeated measurement design with an equivalent control group. This evaluation design will measure the degree to which the CDP's general objective (which was agreed upon by consensus following the EA) is met.

We conclude by recommending to evaluators, agencies and organizations that they integrate EA into their evaluation practices, and publish and disseminate the results obtained, even when they are not positive in terms of program evaluability, such as was the case in this present study, in which the CDP analyzed failed to comply with the requisites for conducting an evaluation of the results in the terms in which the Ibeas de Juarros Local Council wished to do so. Despite this, however, the outcomes and by-products of the EA served to foster learning, bring about improvements and reduce unnecessary costs, as well as, in the medium term, ensure that any future evaluation undertaken will be effective.

## References

- Anney, V.N. (2014). Ensuring the Quality of the Findings of Qualitative Research: Looking at Trustworthiness Criteria. *Journal of Emerging Trends in Educational Research and Policy Studies*, 5(2), 272-281.
- Alvira, F. (1991). *Metodología de la evaluación de programas*. Madrid: CIS.
- Aquilino, N., Arias, E., Estévez, S., & Suaya, A. (2015). Hacia un análisis de evaluabilidad de planes y programas sociales, *Studia Politicae*, (34), 37-72.
- Belford, M., Robertson, T., & Jepson, R. (2017) Using evaluability assessment to assess local community development health programmes, *Medical Research Methodology*, 17, (70). DOI 10.1186/s12874-017-0334-4.
- Bitsch, V. (2005). Qualitative research: A grounded theory example and evaluation criteria. *Journal of Agribusiness*, 23(1), 75-91.
- Craig, P., & Campbell, M. (2015). *Evaluability Assessment: a systematic approach to deciding whether and how to evaluate programmes and policies*. Retrieved from <http://whatworksscotland.ac.uk/wp-content/uploads/2015/07/WWS-Evaluability-Assessment-Working-paper-final-June-2015.pdf>
- Davies, R. (2013). *Planning Evaluability Assessments. A Synthesis of the Literature with Recommendations*. Department for International Development (DFID). Retrieved from [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/248656/wp40-planning-eval-assessments.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/248656/wp40-planning-eval-assessments.pdf).
- Davies, R., & Payne, L. (2015). Evaluability Assessments: Reflections on a review of the literature. *Evaluation*, 21(2), 216-231.
- D'Ostie-Racines, L., Dagenais, C., & Ridde, V. (2013). An Evaluability Assessment of a West Africa Based Non-Governmental Organization's Progressive Evaluation Strategy. *Evaluation and Program Planning*, 36(1), 71-79.

- Fernández- Ballesteros, R. (Ed.). (1995). *Evaluación de programas. Una guía práctica en ámbitos sociales, educativos y de salud*. Madrid: Síntesis.
- Flick, U. (2014). *An introduction to qualitative research* (fourth edition). London: Sage.
- Greene, J. C. (2005). Stakeholders. In S. Mathison (Ed.), *Encyclopedia of Evaluation*. California: Sage.
- Guba, E., & Lincoln, Y. (1989). *Fourth generation evaluation*. California: Sage.
- Houghton, C., Casey, D., Shaw, D., & Murphy, K. (2013). Rigour in qualitative case-study research. *Nurse researcher*, 20(4), 12-17.
- Hurworth, R. (2008). Program clarification: an overview and resources for evaluability assessment, program theory and program logic. *Evaluation Journal of Australasia*, 8(2), 42-48.
- Instituto Navarro de Administración Pública /Navarre Institute of Public Administration (NIPA) (2008). *Protocolo de evaluabilidad de las políticas públicas en la comunidad foral de Navarra*. Pamplona: Gobierno de Navarra.
- Kaufman-Levy, D., & Poulin, M. (2003). *Evaluability assessment: examining the readiness of a program for evaluation: program evaluation briefing series*. Washington: Juvenile Justice Evaluation Center.
- Lincoln Y.S. & Guba, E.G. (1985). *Naturalistic Inquiry*. Newbury Park CA: Sage.
- March, M.X., Orte, C., & Ballester, L. (2016). La pedagogía social en España: de la reconstrucción académica y profesional a la incerteza científica y social. *Pedagogía Social*, 27, 45-82.
- Reiman, C. (2012). *Evaluability Assessments in Peacebuilding Programming*. Reflecting on Peace Practice Program. Retrieved from

[http://dmeforpeace.org/sites/default/files/Wed%20Ready\\_Reimann\\_Evaluability%20Assessment%20Working%20Paper\\_Final\\_20121221.pdf](http://dmeforpeace.org/sites/default/files/Wed%20Ready_Reimann_Evaluability%20Assessment%20Working%20Paper_Final_20121221.pdf)

Ritchie, J. (2003). The Applications of Qualitative Methods to Social Research. In J. Ritchie, J. Lewis, C.M. Nicholls & R. Ormston, R. (Eds.). (2013). *Qualitative research practice: A guide for social science students and researchers*. Sage.

Ruiz-Olabuénaga, J. I. (2012). *Metodología de la investigación cualitativa*. Bilbao: Deusto.

Rutman, L. (1980). *Planning useful evaluations: evaluability assessment*. California: Sage.

Smith, M. F. (1989). *Evaluability assessment: a practical approach*. Boston: Kluwer academic Publisher.

Smith, M. F. (2005). Article on Evaluability assessment. *Encyclopedia of Evaluation*. Thousand Oaks. Sage.

Thurston, W., & Potvin, L. (2003). Evaluability assessment: a tool for incorporating evaluation in social change programs. *Evaluation*, 9(4), 453- 469.

Tobin, G. A., & Begley, C. M. (2004). Methodological rigour within a qualitative framework. *Journal of Advanced Nursing*, 48(4), 388-396.

Trevisan, M. (2007). Evaluability assessment from 1986 to 2006. *American Journal of Evaluation*, 28(3), 290-303.

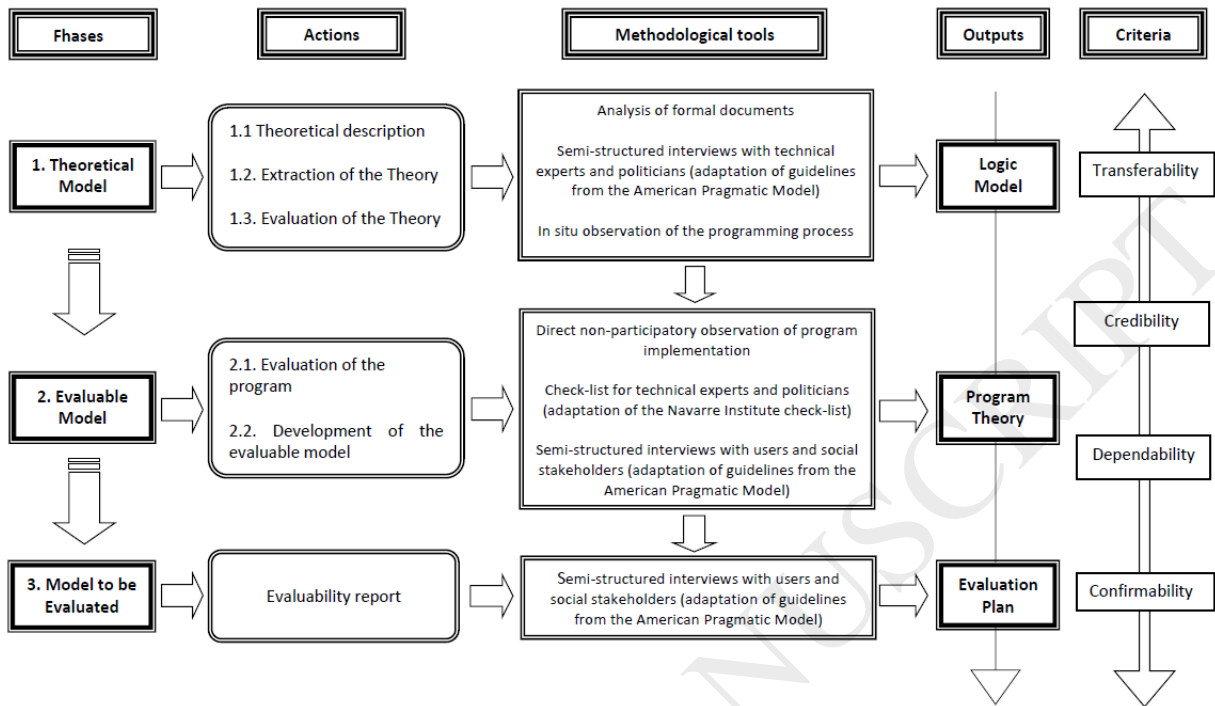
Trevisan, M. S., & Walser, T. M. (2015). *Evaluability Assessment: Improving Evaluation Quality and Use: Improving Evaluation Quality and Use*. California: Sage

Vedung, E. (1997). *Evaluación de políticas públicas y programas*. Madrid: Ministerio de Trabajo y Asuntos Sociales.

Wholey, J.S. (1979). *Evaluation: promise and performance*. Washington, D.C: Urban Institute Press.

ACCEPTED MANUSCRIPT



**Figure 1.** The evaluability assessment model applied in the study.**Figure 2.** CDP flow chart

GOAL: "To improve the subjective quality of life of users through participation in programs, activities, workshops and other initiatives and services which, through the use of instruments inherent to social and cultural animation in recreational activities, serve to enhance their physical, psychological, social and emotional well-being, and to foster social cohesion in the municipality and the integration of all its inhabitants through social engagement."

